(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 0 976 835 A1

(12) EUROPEAN PATENT APPLICATION

(43) Date of publication:
02.02.2000 Bulletin 2000/05

(51) Int. Cl.[7]: **C12Q 1/68**

(21) Application number: **98202549.6**

(22) Date of filing: **29.07.1998**

(84) Designated Contracting States:
AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU MC NL PT SE
Designated Extension States:
AL LT LV MK RO SI

(71) Applicant: **KEYGENE N.V.**
6700 AE Wageningen (NL)

(72) Inventors:
- **Vuylsteke, Marnik Johan Roger**
  6715 EA Ede (NL)
- **Vos, Petrus Antonius Josephina**
  3927 BD Renswoude (NL)
- **Zabeau, Marcus Florent Oscar**
  9000 Gent (BE)

(74) Representative:
de Bruijn, Leendert C. et al
Nederlandsch Octrooibureau
P.O. Box 29720
2502 LS Den Haag (NL)

Remarks:
The applicant has subsequently filed a sequence listing and declared, that it includes no new matter.

(54) Method for detecting nucleic acid methylation using AFLPTM

(57) The invention relates to a method for analyzing of determining the methylation pattern of a starting DNA and/or for distinguishing between methylated and non-methylated sites in the starting DNA, comprising at least

(A) generating a first DNA fingerprint, containing bands corresponding to both the methylated and non-methylated sites of interest; and/or
(B) generating a second DNA fingerprint, containing bands corresponding only to the methylated sites of interest;
and optionally comprising
(C) generating a third DNA fingerprint, containing bands corresponding only to the non-methylated sites of interest;
and optionally further comprising
(D) analysing the fingerprint(s) thus obtained.

Said fingerprints are preferably generated using AFLP, by means of a frequent cutter and a methylation sensitive rare cutter.

The invention further relates to specific methods for generating the above first and second DNA fingerprint by means of AFLP, and kits for use with said methods.

EP 0 976 835 A1

Description

[0001] The present invention relates to a method for detecting DNA methylation using AFLP™. In particular, this method can be used to distinguish between methylated and non-methylated sites (nucleotides) in a nucleic acid, more particular between methylated and non-methylated restriction sites. Thus, the method of the invention can provide information on the methylation pattern of the DNA, which can be visualised as a DNA-fingerprint.

[0002] Selective restriction fragment amplification or AFLP is known, for instance from the European patent application 0 534 858 by applicant, incorporated herein by reference. In general, AFLP comprises the steps of:

(a) digesting a nucleic acid, in particular a DNA, with one or more specific restriction endonucleases, to fragment said DNA into a corresponding series of restriction fragments;
(b) ligating the restriction fragments thus obtained with at least one double-stranded synthetic oligonucleotide adapter, one end of which is compatible with one or both of the ends of the restriction fragments, to thereby produce tagged restriction fragments of the starting DNA;
(c) contacting said tagged restriction fragments under hybridizing conditions with at least one oligonucleotide primer;
(d) amplifying said tagged restriction fragments hybridized with said primers by PCR or a similar technique so as to cause further elongation of the hybridized primers along the restriction fragments of the starting DNA to which said primers hybridized; and
(e) identifying or recovering the amplified or elongated DNA fragment thus obtained.

[0003] The amplified DNA-fragments thus obtained can then be analysed and/or visualised, for instance by means of gel-electrophoresis. This provides a genetic fingerprint showing specific bands corresponding to the restriction fragments which have been linked to the adapter, have been recognized by the primer, and thus have been amplified during the amplification step. The fingerprint thus obtained provides information on the specific restriction site pattern of the starting DNA, and thus on the genetic make-up of the organism from which said DNA has been derived.

[0004] AFLP can therefore be used to identify said DNA; to analyse it for the the presence of specific restriction site patterns, restriction fragment length polymorfisms (RFLP's) and/or specific genetic markers (so-called "AFLP-markers), which may be indicative of the presence of certain genes or genetic traits; or for similar purposes, for instance by comparing the results obtained to DNA-samples of known origin or restriction pattern, or data thereon.

[0005] The primers used in AFLP are such that they recognize the adapter and can serve as a starting point for the polymerase chain reaction. To this end, the primers must have a nucleotide sequence that can hybridize with (at least part of) the nucleotide sequence of the adapter adjacent to the 3' end of the restriction fragment to be amplified. The primers can also contain one or more further bases (called "selective bases") at the 3'-end of their sequence, for hybridization with any complementary base or bases at the 3'-end of the adapter ligated restriction fragment. As, of all the adapter ligated restiction fragments present in the mixture, only those fragments that contain bases complementary to the selective bases will subsequently be amplified, the use of these "selective" primers will reduce the total amount of bands in the final fingerprint, thus making the fingerprint more clear and more specific. Also, the use of different selective primers will generally provide differing fingerprints, which can also be used as a tool for the purposes of identification or analysis.

[0006] As AFLP provides amplification of both strands of a double stranded starting DNA, AFLP advantageously allows for exponential amplification of the fragment, i.e. according to the series 2, 4, 8, 16, etc.. Also, AFLP requires no prior knowledge of the DNA sequence to be analysed, nor prior identification of suitable probes and/or the construction of a gene library from the starting DNA.

[0007] For a further description of AFLP, its advantages, its embodiments, as well as the techniques, enzymes, adapters, primers and further compounds and tools used therein, reference is made to EP-0 534 858, incorporated herein by reference. Also, in the description hereinbelow, the definitions given in paragraph 5.1 of EP-0 534 858 will be used, unless indicated otherwise.

[0008] It is well known that the DNA of a prokaryotic or eukaryotic organism can contain methylated sites, i.e. that certain nucleotides of said DNA strands can be substituted with a methyl-group. In particular, cytosine residues, as well as adenine residues (in bacteria), can be methylated; for instance, in mammals, it is known that 2-7 % of all cytosine-residues are methylated, and this may be as high as 30% in plants. Methylated cytosines can occur as mCG doublets, as small palindromic

5'-^{m}C p G -3'
3'- G p C^{m} -5'

sequences, with both cytosine residues being methylated, or as ^{m}CNG triplets (the latter particular in plants). Often, the

majority of CG-sites in the DNA of both eukaryotes and bacteria are methylated.

[0009] In prokaryotic organisms, the pattern of DNA-methylation can be used to identify a particular bacterial strain or to distinguish replicated and non-replicated DNA (vide B. Lewin, GENES V, Oxford Univ. Press 1994, chapter 20). DNA-methylation also plays a role in DNA repair and the timing of DNA replication.

[0010] In eukaryotes, DNA-methylation is known to be involved in several genetic mechanisms, such as the regulation of gene expression, for instance through gene silencing or gene activation (vide B. Lewin, GENES V, Oxford Univ. Press 1994, chapter 28).

[0011] Also, in eukaryotes, DNA methylation is thought to be associated with genetic diseases through the mechanism of "imprinting", as well as increased susceptibility for mutagenesis and the origin of cancer. For instance, in female individuals, DNA-methylation, which is involved in X-chromosome activation/inactivation, can be used for distinguishing between neoplastic (clonal) cell populations and pseudoplastic or hyperplastic populations, to determine whether a growth of these cells is malignant or not (WO 96/27024). Also, the state of methylation of reporter genes has been used in *in vivo* mutagenicity assays (WO 93/17123 and the references cited therein).

[0012] Furthermore, changes in DNA-methylation can occur during gene transformation, making it possible to distinguish transformed and not-transformed genes or sequences. For instance, analysis of (the changes in) DNA methylation pattern has been used for the early detection of transgenic embryo's (WO 92/22647).

[0013] In plants, such as pea and tomato, methylation patterns can be used to distinguish between varieties by detecting restriction fragment length polymorphisms characteristic of these distinctive varieties (WO 90/05195). In tomato, this is carried out by digesting genomic DNA with a non-methylation sensitive restriction enzyme, and screening the fragments thus obtained by means of Southern hybridization using detectably labelled probes, said probes having been obtained by digestion of tomato genomic DNA with a methylation sensitive restriction enzyme. The bands in the resulting DNA-fingerprint enable identification of species specific, variety specific or individual RFLP's.

[0014] However, this method is not suited or intended for specifically distinguishing between methylated and non-methylated sites within the target DNA. Also, the genomic DNA to be analysed is itself not treated with a methylation sensitive restriction enzyme (these are only used in generating the probes). Furthermore, the technique described in WO 90/05195 does not involve any DNA-amplification step, and suffers from the general disadvantages of similar conventional RFLP detection techniques, such as low resolution, as well as being time consuming and laborious (compare EP-A-0 534 858, paragraph 2.1).

[0015] Nucleotide methylation using sequence specific methylases or restriction methylases has also been used as a tool for marking specific DNA strands or fragments. For instance, in WO 93/22462, thus marked DNA fragments are used in genomic mapping. Of course, when such methylation markers are used, a suitable technique for distinguishing methylated from non-methylated DNA fragments, or even specific methylated and non-methylated sites, is required for subsequent identification and tracking of the marked sequences.

[0016] For a further discussion of DNA methylation in prokaryotes and eukaryotes reference is made to the standard handbooks, to US-A-4,405,760, WO 90/05195 and the further above cited applications, as well as the references cited therein, which are incorporated herein by reference.

[0017] It is known that certain restriction enzymes cleave their target sites according to the state of methylation thereof, for instance only when the target site is non-methylated. Such enzymes, well known examples of which are *Pst*I, *Hpa*II, *Msp*I and *Cla*I, can be used to assess the state and/or degree of methylation of DNA-sequences, or the presence of specific methylated sites therein, by means of selective restriction of (generally) the non-methylated sites, followed by identification of the fragments obtained, usually by means of probing techniques (vide US-A-5,405,760, DE-PS-293 139 and WO 93/22462). However, these methods are cumbersome, inaccurate, and generally require the development of suitable probes and/or at least some prior knowledge of the sequence to be analysed.

[0018] Other restriction enzymes, such as *Mse*I, cut their target site irrespective of whether it is methylated or not. Although both types of restriction enzymes can be and are used in the AFLP technique described in EP-0 534 858 (vide for instance Example 2), this method is as such not suitable determining the methylation pattern of the starting DNA.

[0019] In view of the above, a quick and reliable method for obtaining information on the methylation pattern of a starting or target DNA would be of great value to the art. Therefore, a method for distinguishing between methylated and non-methylated sites in a starting nucleic acid, and in particular for distinguising between methylated and non-methylated restriction sites of genomic DNA, has now been developed. In general, this method comprises at least:

(A) generating a first DNA fingerprint, containing bands corresponding only to the methylated sites of interest; or
(B) generating a second DNA fingerprint, containing bands corresponding to both the methylated and non-methylated sites of interest;
and further optionally comprises
(C) generating a third DNA fingerprint, containing bands corresponding only to the non-methylated sites of interest; as well as
(D) analysing and/or comparing the fingerprint(s) thus obtained;

the respective fingerprints preferably being generated using an amplification technique, such as the AFLP technique described in EP-A-0 534 858.

[0020] By comparing at least two of the fingerprints resulting from (A), (B) and/or (C), information on the methylation pattern of a starting DNA can be derived, as will further be described hereinbelow.

[0021] Preferably, at least the fingerprints resulting from (A) and (B) are generated and compared to each other; more preferably, all three fingerprints resulting from (A), (B) and (C) are generated and compared to each other.

[0022] Conveniently, the above fingerprints can be generated in conjunction with each other, i.e. in parallel and more or less simultaneously, using the same pool or preparation of starting DNA, such as a preparation of intact genomic DNA directly as obtained from the organism of interest. The resulting DNA-fingerprints can then be run in seperate lanes of the same electrophoresis gel, allowing immediate and easy comparison of the resulting patterns of bands. In this way, from one gel, information on the methylation pattern can directly be obtained. Also, each fingerprint can conveniently be generated in a "one pot reaction" using well-established AFLP-technology and equipment.

[0023] However, although for reasons of reliability and easy comparison, it is generally preferred to generate and run the above fingerprints together, preferably starting from the same DNA or DNA-preparation, such as a preparation of intact genomic DNA as isolated from the organism of interest, it is also possible to compare the patterns generated from (1) and/or (2) to known DNA-fingerprints or earlier obtained results, such as a database. This equivalent method is also encompassed within the scope of the present invention. Also, it should be understood that instead of the preferred method of generating a DNA-fingerprint, equivalent methods for analysing and/or visualising restriction fragment mixtures, in particular on the basis of differences in size/molecular weight of the various fragments generated, can also be used, as will be clear to the skilled person.

[0024] Further aspects of the invention reside in the way in which the above indicated DNA fingerprints are obtained, as well as AFLP-methodology used therein. Other aspects reside in the amplified fragments or mixtures of fragements that are obtained according to the methods (A) and (B) described below.

[0025] Yet another aspect comprises any data generated by the method of the invention, optionally on a suitable data carrier, such as paper or a computer disk. This includes the generated DNA-fingerprints (including the gels), photographs or other reproductions thereof, as well as (stored) analog or digital data thereon.

[0026] The invention also comprises kits for use in the invention, comprising at least: a frequent cutter restriction enzyme; a methylation sensitive rare cutter restriction enzyme; (at least) a first and second adapter for use with the frequent cutter; and (at least) a first and second adapter for use with the rare cutter; as well as primers for use with these adapters; wherein these components are essentially as described herein.

[0027] The kits can further contain all known components for AFLP kits, such as restriction enzymes (in which case the adapters are preferably suited to be ligated to the restrictes sites generated with said enzyme); a polymerase for amplification, such as *Taq*-polymerase; nucleotides for use in primer extension; as well as buffers and other solutions and reagentia; manuals, etc.. Further reference is made to the European patent application 0 534 858, incorporated herein by reference.

[0028] Also, it should be noted that the method of the invention comprises and/or combines several features, which as such could also be used in other AFLP-techniques and/or applications, as will be discussed hereinbelow. These form seperate aspects of the invention.

[0029] The method of the invention will be illustrated by means of the enclosed figures, in which

Fig. 1 schematically shows the general method (A) for obtaining a DNA fingerprint showing only the methylated sites; and

Fig. 2 schematically shows the general method (B) for obtaining a DNA fingerprint showing both the methylated and non-methylated sites;

Fig. 3 schematically shows a general method for obtaining a DNA fingerprint containing only the non-methylated sites.

Fig. 4 shows the formation of single strand DNA loop structures in fragments of suitable length that have been ligated to the hybridizing or the same adapter at each end.

Fig. 5 shows the differences in the essential bands obtained in the methods of figures 1, 2 and 3.

Figure 6 gives a general schematic representation of the method of the invention for the enzyme combination (EC) *PstI/MseI*, in which the above methods A, B and C (as shown in figures 1-3) are shown side-by-side.

Figures 7A and B show methylation AFLP fingerprints Figure 2 shows a number of typical methylation AFLP fingerprints obtained with the two step amplification strategy of genomic DNAs from the three maize inbred lines A7, B73 and Mo17.

Figures 8A and 8B are DNA fingerprints showing the segregation of (A) *PstI/MseI* ^{m}AFLP markers and (B) *PstI/MseI* AFLP markers in a Recombinant Inbred (RI) population derived from a cross of two maize lines.

[0030] The starting DNA used in the invention can be any DNA, which contains, is suspected to contain, or is to be

investigated for, DNA-methylation (including hemi-methylated DNA). Usually, native methylated DNA, in particular genomic DNA, will be used; when genomic DNA is used, the method of the invention will generally be used to distinguish between native methylated and native non-methylated sites, in which case method (A) above will be used to generate a first DNA fingerprint, containing bands corresponding only to the native methylated sites of interest; method (B) above will be used to generate a second DNA fingerprint, containing bands corresponding to both the native methylated and native non-methylated sites of interest; and method (C) -if applied- will be used to generate a third DNA fingerprint, containing bands corresponding only to the native non-methylated sites of interest.

[0031] The starting DNA can be derived from any suitable source, such as prokaryotic or eukaryotic organisms (including viruses, yeasts and bacteria), depending upon the intended application. Preferably, eukaryotic DNA, more preferably plant or animal (including human) derived DNA, is used. Also, DNA that has been provided with methylation (i.e. as a marker) or that has been subject to a methylating treatment, can be used.

[0032] Furthermore, instead of a DNA, the invention can be applied to other types of methylated nucleic acids, such as methylated single strand RNA that can occur naturally in the cell.

[0033] In the method of the invention, two different restriction enzymes are used: one enzyme which serves the purpose of reducing the size of the restriction fragments to a range of sizes which are amplified efficiently, hereinbelow referred to as the "frequent cutter", and another enzyme which serves the purpose of targeting rare sequences, hereinbelow referred to as the "rare cutter". For the terms rare cutter and frequent cutter, reference is also made to EP-A-721 987 by applicant, incorporated herein by reference.

[0034] At least one of the enzymes used must be sensitive to (the state of) methylation of the intended site, i.e. be able to distinguish between methylated and non-methylated sites on the target DNA. In practice, generally a methylation-sensitive rare cutter is used.

[0035] Examples of suitable frequent cutter enzymes are *Mse*I and *Taq*I.

[0036] Examples of commercially available suitable methylation-sensitive rare cutters are *Pst*I, *Hpa*II, *Msp*I, *Cla*I, *Hha*I, *EcoR*II, *Bst*BI, *Hin*P1, *Mae*II, *Bbv*I, *Pvu*II, *Xma*I, *Sma*I, *Nci*I, *Ava*I, *Hae*II, *Sal*I, *Xho*I and *Pvu*II, of which *Pst*I, *Hpa*II, *Msp*I, *Cla*I, *EcoR*II, *Bst*BI, *Hin*P1 and *Mae*II are preferred. Other suitable restriction enzymes are for instance described in US-A-5.487,994, US-A-5,340,733, or will be clear to the skilled person. Methylation-senstive mutants of these and other restriction enzymes can also be used.

[0037] For analysis of eukaryotic genomic DNA, preferably C-methylation sensitive restriction enzymes are used. Also, in the practice of the invention, generally restriction enzymes will be used that restrict the non-methylated sites, but not the corresponding methylated sites. However, enzymes which restrict the methylated sites, and not the corresponding non-methylated sites, such as *Dpn*I, can also be used analogously.

[0038] Of the above restriction enzymes, the use of those enzymes which, after digestion, leave the restricted double stranded DNA with a protruding 3'-end, such as *Pst*I, are preferred, for reasons which will be discussed hereinbelow.

[0039] Generally, only one frequent cutter and one rare cutter will be used in generating the fingerprints resulting from (A), (B) and/or (C). However, the use of more than one frequent cutter, and in particular of more than one rare cutter, is also encompassed within the invention. Also, in order to enable comparison of the fingerprints obtained, the restriction enzymes used in generating the fingerprints resulting from (A), (B) and/or (C) should generally be the same, although the use of different enzymes in these methods is encompassed within the scope of the invention, so long as the fingerprints obtained allow comparison to each other or to known data, or otherwise provide information on the methylation of the starting DNA.

[0040] For reasons of convenience, the invention will be further illustrated hereinbelow using the non-limiting combination of the frequent cutter *Mse*I and the methylation sensitive rare cutter *Pst*I. It should however be understood that any other suitable combinations of frequent cutters and rare cutters, for instance chosen from the enzymes mentioned above, can also be used.

[0041] Also, it should be understood that although the procedures (A), (B) and (C) constitute the most convenient and preferred modes of generating the data (fingerprints) on the methylation pattern, equivalent methods or steps, in particular equivalent amplification techniques, can also be used, so long as the obtained results provide information on the methylation of the starting DNA.

[0042] The method of the invention can further be carried out using well known techniques of genetic manipulation, such as described in Sambrook et al, "Molecular Cloning: A Laboratory Manual" (2nd.ed.), Vols. 1-3, Cold Spring Harbor Laboratory (1989) as well as F. Ausubel et al, eds., "Current protocols in molecular biology", Green Publishing and Wiley Interscience, New York (1987), as well as standard AFLP methodology and equipment, which are used in a manner known per se or analogously thereto. For AFLP, reference is again made to EP-A-0 534 858, incorporated herein by reference. The restriction enzymes and other commercially available products are generally used according to the manufacturers guidelines.

A. Method for producing a fingerprint containing bands corresponding to the methylated sites.

[0043] This aspect of the invention comprises:

1. restricting the starting DNA with a frequent cutter and a methylation sensitive rare cutter;
2. ligating the restriction fragments thus obtained to a first and second adapter for the frequent cutter, as well as a first adapter for the rare cutter;
3. amplifying the mixture thus obtained, using primers for the first and second frequent cutter adapter;
4. restricting the amplified fragments thus obtained using a methylation sensitive rare cutter;
5. ligating the restriction fragments thus obtained to a second adapter for the methylation sensitive rare cutter;
6. amplifying the restriction fragments thus obtained using a primer for the first or second frequent cutter adapter, and a primer for the second rare cutter adapter;
7. optionally analysing the mixture of amplified fragments thus obtained in a manner known per se, preferably by generating a DNA fingerprint.

[0044] Method (A) is schematically outlined in Figure 1 and (together with methods B and C) in Figure 5, which both show a starting DNA (generally full length -i.e. uncut- genomic DNA) containing frequent cutter (*MseI*) sites, as well as methylated and non-methylated rare cutter (*PstI*) sites.

[0045] In step 1, the starting DNA is restricted using both the frequent cutter *MseI* and the rare cutter *PstI*, in a manner known per se. This restricts both the *MseI* sites and the non-methylated *PstI*-sites, but leaves the methylated *PstI*-sites intact. Preferably, the *PstI*- and *MseI*-restrictions are carried out simultaneously using a mixture of these enzymes, although it is also possible to use two seperate restriction steps.

[0046] In step 2, the resticted mixture is ligated to adapters, also in a manner known per se, using a mixture of two different *MseI*-adapters, i.e. a "first" and a "second" *MseI* adapter, as well as a first *PstI*-adapter (Said adapters are indicated as "M1", "M2" and "P1" in Figures 1-4, respectively. The terms "first", "second" adapter/primer etc. are only used in this disclosure to denote/distinguish the different adapters/primers used. As such, any adapter/primer which meets the requirements set out below can be used as either the first, the second or any further primer/adapter.)

[0047] The first and second *MseI*-adapter, as well as the first *PstI* adapter are essentially as described in EP-A-0 534 858, or analogous thereto, in that they are suited for use as an adapter in AFLP and that they can be ligated to the cut ends of the *MseI*-fragments or *PstI* fragments respectively. In the method of the invention, the first and second *MseI*-adapter used are further preferably such that they cannot hybridize with one another under the conditions used. This prevents the formation of DNA-loop structures, as shown in figure 4, which can hinder amplification of the fragment. Preferably, the first and second *MseI*-adapters are used about equal amounts. The amount of *PstI* adapter (as compared to the amount of *MseI*) can generally be suitably chosen by the skilled person.

[0048] After ligation of the abovementioned adapters, a mixture of different restriction fragments is obtained. For the purposes of the invention, these can be distinguished in the following types.

- *MseI*/*MseI*-fragments, which either do (indicated as I-A in Fig. 1) or do not (indicated as II-A in Fig.1) contain a non-restricted methylated *PstI*-site. Of these fragments, in Fig.1, only those with the different *MseI*-adapters at each end are shown. Fragments with the same *MseI*-adapter at each end (not shown) will also be obtained, but these will not be amplified efficiently due to the loop formation described above.
- *PstI*/*MseI*-fragments, which contain either the first or second *MseI*-adapter at one end, and the first *PstI*-adapter at the other (of these two, in Fig. 1, only the fragment with the first *MseI*-adapter are shown as III-A/III-B). Rarely, these fragments can also contain a non-restricted methylated *PstI*-site. -
- *PstI*/*PstI*-fragments (shown as IV-A in figure 1). These fragments, which may or may not contain a further methylated *PstI*-site, will be extremely rare, and in general much longer than the *MseI*/*MseI*-fragments I-A/II-A or the *PstI*/*MseI*-fragments III-A, due to the relative abundance of *MseI*-sites in the genome.

[0049] As will be clear to the skilled person, there will generally be several different fragments of each type present in the mixture obtained after adapter ligation, with differing lengths or differing positions of the non-restricted methylated *PstI*-sites, depending upon the restriction site and methylation pattern of the starting DNA. [Also, it should be understood that for each *MseI*/*MseI* fragment shown in Fig. 1, there is a corresponding fragment (not shown) in which the first and second *MseI* adapter are reversed (i.e. bound to the opposite ends of the fragment). For the purposes of the present method, these can essentially be considered equivalent to the fragments shown in Fig. 1, as will become clear hereinbelow.]

[0050] In step 3, the mixture of fragments obtained in step 2 is amplified, using a mixture of primers for the first and second *MseI*-adapter. These primers again are essentially as described in EP-A-0 534 858, or analogous thereto, in that they are suited for use as a primer in AFLP and that they can be hybridize with the first and second *MseI*-adapters

used, respectively. The number of selective nucleotides required in a specific primer/for a specific application may be species-dependant.

[0051] The amplification itself can be carried out in a manner known per se, such as described in EP-A-0 534 858 or in a manner analogous thereto, and is preferably carried out as a (pre)amplification using nonselective primers, i.e. containing no selective nucleotides at the 3'-end of the primer sequence, or at most one selective base. Preferably, the first and second primer are used in about equal amounts.

[0052] During this amplification step, only the *Mse*I-*Mse*I fragments I-A and II-A, which do or do not contain a methylated *Pst*I-site -depending upon the original DNA and the position of the *Mse*I and methylated *Pst*I sites thereon- will be amplified exponentially and effectively. As no *Pst*I-primers are used, the *Pst*I/*Mse*I-fragments III-A will only be amplified linearly, and the *Pst*I/*Pst*I-fragments IV will not be amplified at all.

[0053] Also, as the amplification is carried out in a medium containing only non-methylated nucleotides, the fragments I-A lose their methylation imprint; the corresponding non-methylated fragments as obtained after amplification are indicated as I-A' in Fig.1. These I-A'-fragments can now be restricted using *Pst*I in the subsequent restriction step 4. This restriction will not affect the fragments II-A, which originally did not contain any *Pst*I site.

[0054] Also, in said step 4, as the amplified fragments thus obtained are now non-methylated, instead of a methylation-sensitive restriction enzyme, a methylation insensitive isoschizomer thereof that cuts the same (non-methylated) restriction site can also be used. For instance, when the methylation sensitive *Eco*RII is used, in step 4 its methylation-insensitive isoschizomer *Bst* NI can be used.

[0055] After said restriction step 4, the *Pst*I-cut ends of the restricted fragment I-A' are ligated in step 5 to a second *Pst*I-adapter, the original fragment I-A providing two fragments I-B and I-C, which contain the second *Pst*I-adapter on one end, and either the first or second *Mse*I-adapter on the other. [In fact, as there are originally two equivalent fragments I-A described above, which only differ in the end at which the first and second *Mse*I-adapter, respectively, are bound, also two pairs of equivalent fragments I-B and I-C are obtained (not shown), which have corresponding lengths but only differ in the *Mse*I-adapter opposite to the second *Pst*I-adapter. As will become clear below, for the purposes of the present method, these fragments can be considered equivalent to fragments I-B and I-C, and will not be further distinguished].

[0056] The second *Pst*I adapter (indicated as "P2" in figures 1-4) is again essentially as described in EP-A-0 534 858, or analogous thereto, in that it is suited for use as an adapter in AFLP and that they can be ligated to the *Pst*I cut ends of fragments I-B and I-C. However, the second rare cutter adapter should differ from the first adapter, i.e. not be able to hybridize with the first rare cutter adapter under the conditions used. Also, as mentioned below, the second *Pst*I adapter is preferably the same as the third *Pst*I primer used in step 5 of method B below.

[0057] The respective lengths of fragments I-B and I-C depend upon the length of the original fragment I-A and the position of the methylated *Pst*I-site therein; usually fragments I-B and I-C will differ from each other in length/molecular weight, and will also differ in length/molecular weight from the non-restricted fragments II-A, so that these fragments can be distinguished using a suitable technique.

[0058] To this end, the mixture is amplified in step 6 using suitable primer for the second *Pst*I-adapter, as well as at least one suitable primer for the first and/or second *Mse*I-adapters.

[0059] Most preferably, besides the *Pst*I-primer, only one *Mse*I-primer (i.e. for either the first or the second *Mse*I-adapter) as used. In this way, only the fragments containing a *Pst*I-adapter at one end, and the corresponding *Mse*I-adapter at the other end will be amplified exponentially. Other fragments present in the mixture will only be amplified linearly and/or less efficiently, or even not at all, either due to loop formation or because only one or none of the primers required for exponential amplification is available, which will reduce both the total number of bands in the final fingerprint, as well as the number of non-informative bands.

[0060] The amplification step itself can be carried out in a manner known per se, such as described in EP-A-0 534 858 or in a manner analogous thereto, and is preferably carried out as a two-step amplification, the first of which is a (pre)amplification using a primer for the second *Pst*I-adapter with one selective nucleotide at the 3'-end of the primer sequence and a primer for the first (resp. second) *Mse*I-adapter with one selective nucleotide at the 3'-end of the primer sequence (so-called +1 primers). This is followed by amplification using selective primers for the second *Pst*I- and first (resp. second) *Mse*I- adapter, for instance using primers containing three selective nucleotides at the 3'-end (+3 primers). In both the pre-amplification and the final amplification, the respective primers can be used in about equal amounts.

[0061] In step 6, the number of selective nucleotides required in a specific primer for a specific application may be species-dependant. For instance, in maize, +2/+3 amplification and the double +3/+3 selectivity for the *Pst*I/*Mse*I primer combination (PC) may be used.

[0062] The resulting amplified fragments are then visualized using a suitable technique, such as the formation of a DNA fingerprint through gel electrophoresis. This fingerprint will show at least bands corresponding to fragments of types I-B and I-C, which through comparison with fingerprints generated according to method (B) or (C) or with known data can provide information on the methylaton pattern of the starting DNA, as further discussed below.

[0063] The fingerprint can again contain other minor and/or less informative bands, which -although they can still provide some useful information on the starting DNA- are not critical for the purposes of the invention. These bands will generally not interfere with the information essential for the purposes of the invention.

B. Method for producing a fingerprint containing bands corresponding to both the methylated and the non-methylated sites.

[0064] This aspect of the invention comprises:

1. restricting the starting DNA with a frequent cutter;
2. ligating the restriction fragments thus obtained with a third and fourth adapter for the frequent cutter;
3. amplifying the mixture thus obtained, using primers for the third and fourth adapter for the frequent cutter;
4. restricting the amplified fragments thus obtained using a methylation-sensitive rare cutter;
5. ligating the restriction fragments thus obtained to a third adapter for the methylation sensitive rare cutter;
6. amplifying the mixture thus obtained, using a primer for the third methylation sensitive rare cutter adapter, as well as a primer for the third or fourth frequent cutter adapter;
7. optionally analysing the mixture of amplified fragments thus obtained in a manner known per se, preferably by generating a DNA fingerprint.

[0065] Method (B) is schematically outlined in Figure 2, which shows a DNA fragment containing frequent cutter (*MseI*) sites (again indicated as minor arrows in the sequence), as well as methylated (large arrow with cross) and non-methylated (large arrow) rare cutter (*PstI*) sites.

[0066] In method (B), the third and fourth frequent cutter adapter and the second and third rare cutter adapter are again essentially as described in EP-A-0 534 858, or analogous thereto, in that they are suited for use as an adapter in AFLP and that they can be ligated to the cut ends of the frequent cutter and rare cutter fragments, respectively. The two frequent cutter adapters are preferably again such that they cannot hybridize with one another under the conditions used, in order to prevent undesired loop formation.

[0067] The third and fourth cutter adapters can differ in sequence and/or number of nucleotides from the first and second frequent cutter adapter used in method (A) above, or method (C) below. Preferably, however, for reasons of convenience and in order to minimize the total amount of different adapters required in carrying out the invention, the third and fourth adapter for the frequent cutter restriction enzyme are the same as the first and second frequent cutter adapter used above, respectively. This also allows an even more direct comparison of the fingerprints obtained, thus making the entire method even more reliable.

[0068] Also, although the third rare cutter adapter can differ in both sequence and number of nucleotides from the first and second rare cutter adapter used in method (A), preferably the third frequent cutter adapter used in method (B) is the same as the second rare cutter adapter used in step 5/6 of method (A) above.

[0069] Hereinbelow, in accordance with these preferred embodiment, the third and fourth frequent cutter adapter and the third rare cutter adapter will be denoted hereinbelow as the "first and second" *MseI* primers and the "second" *PstI*-primer.

[0070] In step 1, the starting DNA is restricted using the frequent cutter *MseI*, in a manner known per se. This restricts the *MseI*-sites but leaves both the methylated and non-methylated *PstI*-sites intact.

[0071] In step 2, the restricted mixture is ligated to adapters, also in a manner known per se, using a mixture of the first and second *MseI*-adapters.

[0072] After ligation of the frequent cutter adapters, a mixture of different restriction fragments is obtained. Of these fragments, in Fig. 2, only the fragments containing different *MseI* adapters at each end are shown. It should however be understood that corresponding fragments (not shown) containing either the first or the second frequent cutter adapter at both ends of the fragment are also obtained. Due to the loop formation mentioned above, these fragments will not be amplified efficiently during the subsequent amplification step.

[0073] Of the fragments containing the first and the second frequent cutter adapter shown in figure 2, for the purposes of the present method, three different types of fragments can be distinguished, i.e. one type containing the methylated *PstI*-site(s) (indicated as V-A in Fig.2), a second type containing the originally non-methylated *PstI*-site(s) (indicated as VI-A in Fig.2), and a third type containing no *PstI*-sites (indicated as VII in Fig.2).

[0074] As above, there will again be several different fragments of each type present in the mixture obtained after adapter ligation, with differing lengths and/or (in fragments V-A and VI-A) position of the *PstI*-sites, depending upon the restriction site pattern and methylation pattern of the starting DNA. [Also, it should again be understood that for each fragment type shown in Fig. 2, there is a corresponding fragment (not shown) in which the first and second adapter are reversed (i.e. bound to the opposite ends as shown in Fig. 2). For the purposes of the present method, these can essentially be considered equivalent to the fragments shown in Fig. 2, as will become clear hereinbelow.]

[0075] In step 3, the mixture of fragments obtained in step 2 is amplified, using a mixture of primers for the first and second *Mse*I-adapter. These primers again are essentially as described in EP-A-0 534 858, or analogous thereto, in that they are suited for use as a primer in AFLP and that they can hybridize with the first and second *Mse*I-adapters used, respectively. Again, the number of selective nucleotides required in a specific primer/for a specific application may be species-dependant.

[0076] Also, it should be noted that for methods (A), (B) and (C), where possible the primers used in conjunction with each of the respective adapters are preferably essentially the same for each of the adapters used, in that they have the same sequence in the region for hybridization with the adapter, and may only differ in the presence of the selective bases at their 3'-end, i.e. the number thereof and/or the specific nucleotides. This enables the same primers to be used, or the primers to be synthesised in conjuction wit each other and/or starting from the corresponding non-selective primer sequence, so that a limited number of starting materials (primers) are required for carrying out the three methods of the invention.

[0077] The amplification itself can be carried out in a manner known per se, such as described in EP-A-0 534 858 or in a manner analogous thereto, and is preferably carried out as a (pre)amplification using nonselective primers, i.e. containing no selective nucleotides at the 3'-end of the primer sequence, or at most one selective base. Preferably, the first and second primer are used in about equal amounts.

[0078] During the amplification step, exponential amplification of all fragments containing two different *Mse*I-adapters is obtained. Due to the loop-formation described above, fragments containing the same *Mse*I-adapter at each end are not, or less efficiently amplified, and thus are present in minor amounts in the amplified mixture, which do not interfere with the remainder of the procedure.

[0079] Also, as the amplification is carried out in a medium containing only non-methylated nucleotides, the methylated fragments V-A lose their methylation imprint during amplification; the corresponding non-methylated fragments as obtained after amplification are indicated as V-A' in Fig.2. These V-A'-fragments can now be restricted using *Pst*I in the subsequent restriction step 4, together with the amplified fragments VI-A containing the originally non-methylated *Pst*I-sites.

[0080] Again, in said step 4 of method (A) above, instead of a methylation-sensitive restriction enzyme, a methylation insensitive isoschizomer thereof that cuts the same (non-methylated) restriction site can also be used.

[0081] After said restriction step 4, the *Pst*I-cut ends of the restricted fragments are ligated to the third *Pst*I-adapter, again providing a mixture of different adapter-containing restriction fragments, as shown in Fig.2. The third *Pst*I-adapter is again essentially as described in EP-A-0 534 858, or analogous thereto, in that it is suited for use as an adapter in AFLP and that it can be ligated to the cut ends of the *Pst*I-fragments. Also, as mentioned, the third *Pst*I adapter is preferably the same as the second *Pst*I primer used in step 5 of method A above, and is so indicated below.

[0082] After ligation, the fragment(s) V-A/V-A' will essentially provide two restriction fragments (V-B and V-C), each containing the second *Pst*I-adapter at one end, and either the first or second *Mse*I-adapter on the other end. These fragments will differ in length dependent on the position of the methylated *Pst*I-site in the original *Mse*I fragment V-A. [Again, as with fragments I-A and I-B/I-C in method (A) above, it should be understood that for each the fragments V-B and V-C shown in figure 2 (i.e. containing either the first or second *Mse*I adapter), there will be two corresponding fragments (not shown) which contain the opposite *Mse*I-adapter (i.e. containing the second or first *Mse*I adapter, respectively). Of these four fragments, generally only the two fragments containing the first (or the second) *Mse*I-adapter will be amplified in step 6 below.]

[0083] Correspondingly, the fragment(s) VI-A, which originally contained the non-methylated *Pst*I-site, also two different fragments (VI-B and VI-C) are obtained, with different lengths depending on the position of the non-methylated *Pst*I-site in the origina/ *Mse*I-fragment VI-A. [As with fragment V-A and V-B/V-C, in practice two pairs of equivalent fragments with the same length, but with either the first or the second *Mse*I-adapter opposite the *Pst*I-adapter, are obtained, which again are essentially equivalent and will not be distinguished below].

[0084] The mixture obtained after restriction with *Pst*I and ligation of the *Pst*I-adapter will further contain the *Mse*I fragments VII that originally did not contain either a methylated or non-methylated *Pst*I-site.

[0085] In principle, all fragments V-B, V-C, VI-A, VI-B, as well as the fragments VII will have differing lengths/molecular weights, so that they can be distinguished using a suitable technique.

[0086] To this end, the mixture is amplified in step 6 using suitable primer for the second *Pst*I-adapter, as well as at least one suitable primer for the first and/or second *Mse*I-adapters. These primers are again essentially as described in EP-A-0 534 858, or analogous thereto, in that they are suited for use as a primer in AFLP and that they can be hybridize with the second *Pst*I-adapter, and the first and/or second *Mse*I-adapter, respectively. Preferably, selective primers are used.

[0087] As mentioned above, these primers are preferably the same as the corresponding primers used in step 6 of method A above. Also, for the reasons given in method A, most preferably only one *Mse*I-primer (i.e. for either the first or the second *Mse*I-adapter) is used, besides the primer for the first *Pst*I-adapter. Also, it should be noted that, as no primers for the second *Pst*I-adapter are used in step 6, the *Pst*I/*Mse*I-fragments VI generated in the first restriction step

of method B will again only be amplified linearly.

[0088] The amplification step itself can be carried out in a manner known per se, such as described in EP-A-0 534 858 or in a manner analogous thereto, and is preferably carried out as the two-step amplification of step 6 of method A above, i.e. as a +1/+1 (pre)amplification (*PstI*+1/*MseI*+1), followed by a +2/+3 or +3/+3 amplification (*PstI*+2/*MseI*+3 or (*PstI*+3/*MseI*+3). In both the pre-amplification and the final amplification, the respective primers can be used in about equal amounts.

[0089] Again, as in step 6 of method (A), the number of selective nucleotides required in a specific primer for a specific application may be species-dependant. For instance, in maize, +2/+3 amplification and the double +2/+3 selectivity for the *PstI*/*MseI* primer combination (PC) may be used.

[0090] The resulting amplified fragments are then visualized using a suitable technique, such as the formation of a DNA fingerprint through gel electrophoresis. This fingerprint will show at least bands corresponding to fragments of types V-B, V-C, VI-B and VI-C, which though comparison with fingerprints generated according to method (A) or (C) or with known data can provide information on the methylaton pattern of the starting DNA, as discussed below.

[0091] The fingerprint can also contain other minor and/or less informative bands, which -although they can still provide some useful information on the starting DNA- are not critical for the purposes of the invention. These bands will generally not interfere with the information essential for the purposes of the invention.

C. Method for producing a fingerprint containing bands corresponding only to the non-methylated sites.

[0092] This method essentially corresponds to the conventional AFLP-technique described in European patent application 0 534 858, and is therefore as such not a seperate aspect of the invention. However, the use of this method in the context of the present invention, in particular for generating a reference for comparison to the fingerprints obtained from in the above methods A and/or B, is encompassed by the present application. Method (C) comprises:

1. restricting the starting DNA with a frequent cutter and a methylation sensitive rare cutter;
2. ligating the restriction fragments thus obtained to a fifth adapter for the frequent cutter, as well as a fourth adapter for the methylation sensitive restriction;
3. amplifying the mixture thus obtained, using a primer for the fifth frequent cutter adapter, as well as a primer for the fourth methylation sensitive rare cutter adapter;
4. optionally visualising the amplified fragments thus obtained, preferably by means of a DNA fingerprint.

[0093] Method (C) is schematically outlined in Figure 3, which shows a DNA fragment containing frequent cutter (*MseI*) sites (again indicated as minor arrows in the sequence), as well as methylated (large arrow with cross) and non-methylated (large arrow) rare cutter (*PstI*) sites.

[0094] The fifth frequent cutter and the fourth rare cutter adapters are again essentially as described in EP-A-0 534 858, or analogous thereto, in that they are suited for use as an adapter in AFLP and that they can be ligated to the cut ends of the restriction fragments (the frequent cutter and rare cutter fragments, respectively).

[0095] Again, although these adapters can differ in sequence and/or number of nucleotides from the frequent cutter adapters and rare cutter adapters used in method (A) and (B) above, preferably, for the reasons of convenience discussed above, the fifth adapter for the frequent cutter restriction enzyme is preferably the same as the first or the second -preferably the first- frequent cutter adapter used above, and the fourth rare cutter adapters is the same as the first or second - preferably the second- rare cutter adapter used above. In accordance with this preferred embodiment, the fifth frueqent cutter adapter and the fourth rare cutter adapter will be denoted hereinbelow as the "first" *MseI* adapter and the "second" *PstI* adapter.

[0096] Also, the primers used in conjunction with each of the respective adapters are again preferably essentially the same as the corresponding primers used in the methods (A) and (B) above.

[0097] In step 1, the starting DNA is restricted using both the frequent cutter *MseI* and the rare cutter *PstI*, in a manner known per se. This restricts both the *MseI* sites and the non-methylated *PstI*-sites, but leaves the methylated *PstI*-sites intact.

[0098] Again the *PstI*- and *MseI*-restrictions are preferably carried out simultaneously using a mixture of these enzymes, although it is also possible to use two seperate steps.

[0099] In step 2, the resticted mixture is ligated to *MseI*- and *PstI*-adapters, also in a manner known per se. However, besides the second *PstI*-adapter, and different from methods A and B above, only one *MseI*-adapter is used (i.e. either the first or the second), instead of a mixture of two different *MseI*-adapters.

[0100] After ligation of the frequent cutter adapters, a mixture of different restriction fragments is obtained. For the purposes of the invention, these can be distinguished in *MseI*/*MseI*-fragments (indicated as VII in Fig.3), *PstI*/*MseI*-fragments (indicated as VIII-A and VIII-A' in Fig.3), and *PstI*/*PstI*-fragments (indicated as IX in Fig. 3). All these fragments may or may not contain the non-restricted methylated *PstI*-sites (only shown for the *MseI*/*MseI*-fragments, as

VII-A and VII-A' respectively); as this method is only directed to detecting the non-methylated *Pst*I-sites, these sites are not relevant, nor do they interfere with the resulting fingerprint.

[0101] Again, there will be several different fragments of each type present in the mixture obtained after adapter ligation, with differing lengths, depending upon restriction site pattern of the starting DNA. The respective lengths of the fragments VII, VIII and IX will depend on the positions of the *Mse*I-sites and the non-methylated *Pst*I-sites in the starting DNA (but not the methylated *Pst*I-sites). Usually, fragments VII, VIII and IX will differ in length/molecular weight, with the *Pst*I/*Pst*I fragments being appreciably longer than the other fragments due to the relative abundance of the *Mse*I-sites. Because of their differing lengths/molecular weight, these fragments can distinguished using a suitable technique.

[0102] To this end, the mixture is amplified in step 3 using suitable primer for the second *Pst*I-adapter, as well a suitable primer for the first *Mse*I-adapters. As mentioned above, these primers are preferably the same as the corresponding primers used in step 6 of method A and/or B above.

[0103] The amplification step itself can be carried out in a manner known per se, such as described in EP-A-0 534 858 or in a manner analogous thereto, and is preferably carried out as the two-step amplification of step 6 of method A above, i.e. as a (pre)amplification using a +1-primers for the second *Pst*I-adapter and the first *Mse*I-adapter, followed by selective amplification using +3-primers. In both the pre-amplification and the final amplification, the respective primers can be used in about equal amounts.

[0104] Again, the number of selective nucleotides required in a specific primer for a specific application may again be species-dependant. For instance, in maize, +2/+3 amplification and the double +3/+3 selectivity for the *Pst*I/*Mse*I primer combination (PC) may be used.

[0105] It should be noted that during said amplification, the *Mse*I/*Mse*I-fragments VII-A and VII-A' (i.e. with or without a methylated *Pst*I-site) will not be amplified efficiently due to the formation of loop structures. Also, any *Pst*I/*Pst*I-fragments will generally not be amplified efficiently, as they are to long to be amplified for their full length during the cycle times used.

[0106] The resulting amplified fragments are then visualized using a suitable technique, such as the formation of a DNA fingerprint through gel electrophoresis. This fingerprint will show at least bands corresponding to fragments of types VIII-A and and VIII-B, which can be used as a comparison for the fingerprints generated according to method (A) or (B), as discussed below.

D. Determining the methylation pattern of the starting DNA from the fingerprints generated in methods (A), (B) and (C).

[0107] The differing fragments, essential for determining the methylation pattern of the starting DNA, that are generated, amplified and detected in each of the methods (A), (B) and (C), are schematically indicated in Figure 5, which shows a DNA fragment containing frequent cutter (*Mse*I) sites, as well as methylated and non-methylated rare cutter (*Pst*I) sites.

[0108] As can be seen from figure 5, method A will provide fragments, and bands in the DNA-fingerprint, corresponding to the methylated (I-B and I-C) sites, but not to the non-methylated (III-A and III-A') sites.

[0109] Method B will provide fragments, and bands in the DNA-fingerprint, corresponding to the methylated (V-B and V-C) and non-methylated (VI-B and VI-C) sites.

[0110] Method C will provide fragments, and bands in the DNA-fingerprint, corresponding to the non-methylated (VIII-B and VIII-C) sites, but not the methylated (VII-A) sites.

[0111] Also, as will immediately be clear from figure 5, the respective correseponding fragments of interest (i.e. I-B/I-C and V-B/V-C on the one hand, and VI-B/VI-C and VIII-A and VIII-B on the other) will have the same size/molecular weight, and therefore give corresponding bands in the fingerprint, if the respective electrophoreses are run in parallel lanes of the same gel.

[0112] Therefore, comparing the fingerprints generated in method (B) to the fingerprint generated in method (A) will provide information on the originally non-methylated sites, which will occur as bands in fingerprint (B) not present in fingerprint (A) (i.e. bands V-A and V-B). Also, these bands will be present in fingerprint (C) (i.e. bands VIII-A and VIII-B).

[0113] Information on the originally methylated sites can be derived from comparing either fingerprint (A), fingerprint (B), and preferably both, to fingerprint (C): these will occur as bands present in (A) and (B) (i.e. bands I-B/I-C and VI-B/VI-C), but not in (C).

[0114] Also, using well known techniques for analysing the results from DNA-fingerprints derived from a segregating population, further information can de derived on the position of the methylated and non-methylated sites in the genome, for instance on the relative position of the methylated site to known AFLP-markers or other genetic markers in the starting DNA.

[0115] Also, from the restriction site recognised by the methylation sensitive rare cutter used, information can be obtained on the methylated sequence; the use of different methylation sensitive rare cutters providing information on the state of methylation of different sites/sequences in the genome.

[0116] The method of the invention can be used as a general tool for detecting DNA-methylation or methylation patterns, both in prokaryotes and eukaryotes, including viruses, yeasts, fungi, bacteria, plants, animals and humans. As such, it can be used to replace known techniques for determining and/or estimating the extend of DNA-methylation, in all applications in which such information is of interest, such as those discussed above.

[0117] The method of the invention is particularly suited for applications in which a speedy, qualitative determination of the state of methylation or the methylation pattern, that can be applied to large scale testing, is required. Generally, these will be applications similar to those in which conventional AFLP-techniques are also the technique of choice in determining genetic make-up or RFLP's.

[0118] Some non-limiting applications of specific interest are:

- distinguishing, identifying or classifying individuals, varieties or species on the basis of their DNA-methylation, by comparing methylation patterns of different individual sources;
- fingerprinting of genomes and detecting methylation polymorphisms;
- detecting specific methylation patterns corresponding to the presence of specific sequences or genetic traits, as well as inherited DNA-methylation (patterns) and allele-specific methylation (patterns);
- estimating the extent of methylation in a DNA or genome;
- distinguishing transformed and not-transformed genes or sequences, for instance in the detection of transgenes; and distinguishing replicated and non-replicated DNA in bacteria;
- detecting gene silencing or gene activation, and generally investigating the regulation of gene expression, in eukaryotes;
- detecting DNA-methylation associated with genetic diseases, genomic imprinting and cancer;
- determing the methylation of reporter genes in *in vivo* mutagenicity assays;
- detecting methylated DNA-markers;
- detecting site-specific methylation (enzymatically or chemically), methylation sensitive sites and/or mutation-associated methylation.

[0119] The invention will now be illustrated by means of the following non-limiting Examples.

Examples.

[0120] The Examples below describe the novel PCR-based method of the invention for detecting methylation of restriction sites randomly over the genome. Said method will hereinbelow be referred to as "methylation AFLP™".

[0121] The technique is based on (i) the use of a pair of restriction enzymes consisting of a methylation-sensitive rare cutter and a methylation-insensitive frequent cutter, (ii) the comparison of fragments obtained from native and PCR amplified DNA, and (iii) selective amplification of genomic restriction fragments using PCR.

[0122] The power of methylation AFLP resides in its positive display of the unmethylated and the native methylated sites jointly and seperately, as is obtained after selective amplification of the restriction fragments by PCR. The technique is based on selective amplification of genomic restriction fragments obtained directly from native DNA and also from nonselective amplified DNA by using a methylation-sensitive rare cutter and a methylation-insensitive frequent cutter. Nonselective amplification of restriction fragments is achieved by using only the adapter and restriction site sequence as target sites for primer annealing. Selective amplification is achieved by the use of primers that extend into the restriction fragments, amplifying only those fragments with nucleotides flanking the restriction sites that match the primer extensions.

[0123] The advantages of methylation AFLP are manifold: no prior sequence knowledge is needed, a limited set of generic primers is used, a high multiplex ratio (said ratio being a function of the selected specific primer sets) is obtained and a positive display of the unmethylated and native methylated sites is provided. Also, the detection of DNA-methylation is genome-wide. Typically 50-100 restriction fragments are co-amplified and resolved on denaturing polyacrylamide gels. The Examples below will further illustrate how methylation AFLP can be used to estimate the extent of CpG and CpNpG methylation, to detect epialleles and additional sequence polymorphism, and to follow the inheritance of C-methylation.

[0124] A variety of techniques to assess the degree of DNA methylation is presently available, which can be divided into sequence-unspecific and sequence-specific methods. The first category can be used to analyse the different types of modified bases and to quantify them, but do not provide any information about the precise location of the modified site within a given nucleic acid sequence. Into this category fall immunological, chromatographic, electrophoretic and spectrophotometric procedures that follow a complete chemical or enzymatic hydrolysis of the target DNA. Another sequence-independent approach involves the use of methylation-sensitive restriction endonucleases where genomic DNA digests obtained with these restriction enzymes are compared using gel electrophoresis (Saluz, H.P. and Jost, J-P. (1993) Jost, P. and Saluz, H.P.(ed.), *DNA Methylation: Molecular Biology and Biological Significance,* Saluz

Birkhauser Verlag, Basel, Switzerland). Comparison of fragment lengths and intensities of the resulting digestion patterns on the gel, allows estimation of the proportion of methylated sites.

[0125] The second category enables analysis of the precise location of methylated bases within a known DNA sequence. Into this category fall techniques based on the use of pairs of isoschizomeric restriction enzymes that differ in sensitivity to methylation, in combination with Southern-blot analysis (Bird, A.P. and Southern, E.M. (1978), *J. Mol. Biol.* **118**, 27-47; Waalwijk, C. and Flavell, R.A. (1978) *Nucl. Acids Res.* **5**, 3231-3236) or PCR.

[0126] Other PCR-based methods for detecting DNA-methylation have also been reported; one of these PCR-based methods takes advantage of the fact that PCR amplification occurs only if the DNA between the two primer sites remains uncleaved by the methylation-sensitive restriction enzyme *Hpa*II (Singer-Sam, J., LeBon, J.M., Tanguay, R.L. and Riggs A.D. (1990) *Nucl. Acids Res.* **18**, 687; Singer-Sam, J., Grant, M., LeBon, J.M., Okuyama, K., Chapman, V., Monk, M. and Riggs, A.D. (1990) *Mol. Cell. Biol.* **10**, 4987-4989; Heiskanen M., Syvanen, A.C., Siitari, H., Laine, S. and Palotie, A. (1994) *PCR Methods and Applic. 4, 26-30*). Other PCR-based methods combine PCR with genomic sequencing to identify methylated cytosine residues (Maxam, A.M. and Gilbert, W. (1980) *Methods Enzymol.* **65**, 499-560; Pfeiffer, G.P., Steigerwald, S.D., Mueller, P.R., Wold, B. and Riggs A.D. (1989) *Science* **246**, 810-813), utilizing the Maxam and Gilbert chemical cleavage reactions carried out on genomic DNA with various additional procedures to enhance the signal from the sequence under investigation.

[0127] The bisulfite genomic sequencing technique described by Frommer et *al.* (Frommer, M., McDonald L.E., Millar, D.S., Collis, C.M., Watt, F., Grigg, G.W., Molloy, P.L. and Paul C.L. (1992) *Proc. Natl. Acad. Sci.* **89**, 1827-1831) circumvents the drawbacks of the latter methods by eliminating the chemical cleavage reactions and providing a positive identification of 5-methylcytosine residues. Additional methods have been developed recently which utilize bisulfite treatment of DNA as a starting point for methylation analysis. These include methylation-specific PCR (MSP) (Herman, J.G., Graff, J.R., Myohanen, S., Nelkin, B.D. and Baylin, S.B. (1996) *Proc. Natl. Acad. Sci.,* **93**, 9821-9826), restriction enzyme digestion of PCR products amplified from bisulfite-converted DNA (Sadri, R. and Hornsby, P.J. (1996) *Nucleic Acids Res.* **24**, 5058-5059; Xiong, Z. and Laird, P.W. (1997) *Nucleic Acids Res.* **25**, 2532-2534) and bisulfite treatment of DNA followed by single nucleotide primer extension (Gonzalgo, M.L. and Jones, P.A. (1997) *Nucleic Acids Res.* **25**, 2529-2531). With the method of the invention, in contrast with the bisulfite genomic sequencing technique and relatives, the methylation of a virtually unlimited number of loci can be detected.

[0128] The principle of the method of the invention is schematically shown in Figure 6, which gives a schematic representation of the methylation AFLP technique for the enzyme combination (EC) *Pst*I/*Mse*I: 'nonselective amplification' stands for the nonselective *Mse*I/*Mse*I$^+$ amplification; 'selective amplification' stands for the selective *Pst*I/*Mse*I amplification.

[0129] The left hand column shows the generation of a subset of templates (hereinbelow: "subset A" or "A-templates") in accordance with the general method "A" above, providing restriction fragments with a native methylated *Pst*I site, followed by the selective *Pst*I/*Mse*I amplification procedure.

[0130] The middle column shows the generation of a subset of templates (hereinbelow: "subset B" or "B-templates") in accordance with the general method "B" above, providing restriction fragments with a native methylated or unmethylated *Pst*I site, followed by the selective *Pst*I/*Mse*I amplification procedure.

[0131] The right hand column shows the generation of a subset of templates (hereinbelow: "subset C" or "C-templates") in accordance with the general method "C" above, representing the restriction fragments with a native unmethylated *Pst*I site, followed by the selective *Pst*I/*Mse*I amplification procedure.

[0132] The methylation AFLP technique of the invention is based on the amplification by PCR of three derived subsets (A, B, C) of genomic restriction fragments, differing in (i) digestion of the genomic DNA using two restriction enzymes and (ii) ligated double stranded (ds) adapters.

[0133] The general outline of this method is depicted in Figure 6 for the EC *Pst*I/*Mse*I. Subset A, representing only the restriction fragments with a native methylated *Pst*I cutter site, is obtained in the following way (vide Figure 6): the genomic DNA is digested to completion with the two restriction enzymes *Pst*I and *Mse*I, and the corresponding ds (AFLP) *Pst*I-adapter$^+$ and MseI-adapters$^{(+)}$ are ligated to the unmethylated *Pst*I sites and *Mse*I sites to block the unmethylated *Pst*I ends and to generate template DNA for the nonselective *Mse*I/*Mse*I$^+$ amplification. Two different *Mse*I-adapters are chosen to avoid stem-loop structure forming of the *Mse*I-*Mse*I fragments, which have an inverted repeat at the ends. This is followed by the *Mse*I/*Mse*I$^+$ nonselective amplification of the genomic restriction fragments that may carry an internal methylated *Pst*I site. During this PCR-step, restriction sites lose their methylation imprint. A second digestion of the 'demethylated' *Pst*I sites followed by ligation of the *Pst*I adapter generates template DNA for further selective *Pst*I/*Mse*I amplification.

[0134] The subset B-templates, representing the restriction fragments with native methylated or unmethylated *Pst*I sites, are obtained in a similar way as the subset A-templates, with the exception that the genomic DNA is digested to completion with only *Mse*I, and corresponding ds (AFLP) MseI-adapters$^{(+)}$ are ligated to generate template DNA for the nonselective *Mse*I/*Mse*I$^+$ amplification (see Figure 6).

[0135] The subset C-templates, representing only the restriction fragments with a native unmethylated *Pst*I site are

obtained according to the published AFLP procedure (Vos, P., Hogers, R., Bleeker, M., Reijans, M., van de Lee, T., Hornes, M., Frijters, A., Pot, J., Peleman, J., Kuiper, M. and Zabeau, M. (1995) *Nucl. Acids Res.* **23,** 4407-4414), in the manner shown in Figure 6: digestion to completion of the genomic DNA by the EC *PstI/MseI*, and ligation of the corresponding ds AFLP adapters to the unmethylated *PstI* sites and *MseI* sites to generate template DNA for the selective *PstI/MseI* amplification.

[0136] Finally, the A, B and C subsets of templates are simultaneously selectively amplified following the two-step amplification procedure strategy (pre-amplification followed by the AFLP reaction), according to the AFLP fingerprinting method of complex genomes (Vos, P., Hogers, R., Bleeker, M., Reijans, M., van de Lee, T., Hornes, M., Frijters, A., Pot, J., Peleman, J., Kuiper, M. and Zabeau, M. (1995) *Nucl. Acids Res.* **23,** 4407-4414.) and then displayed together in adjacent lanes A, B and C of a methylation AFLP fingerprint.

[0137] The Examples below are carried out in accordance with the preferred embodiments described above, in that as few different adapters and primers as possible are used. Accordingly, in total only two *MseI* adapters (i.e. those indicated above as the "first" and "second" *MseI* adapter), and their two corresponding *MseI* primers (or primer combinations), and two *PstI* adapters (i.e. those indicated above as the "first" and "second" *PstI* adapter) and their two corresponding *PstI* primers (or primer combinations) are used. In the Examples and Figures, the second *MseI* adapter is indicated with a (+), and the first *PstI* adapter is indicated using an (*), i.e. as *MseI*+ and *PstI*+ respectively. Accordingly, the corresponding second *MseI* adapter and the corresponding first *PstI* primer are also indicated with a (+) and a (*), respectively.

Example I: General description and protocols followed.

A. Genomic DNAs and enzymes used.

[0138] The sources of the genomic DNAs and enzymes used were as follows: Tomato DNA (Near Isogenic Lines (NILs) 83M-S and 83M-R obtained from De Ruiter zonen, The Netherlands; cv. Motelle and Mobox obtained from INRA, Montfavet, France; inbred line RC10 obtained from Enza Zaden, The Netherlands; inbred line 52201 obtained from Rijk Zwaan, The Netherlands), maize DNA (inbred lines B73, Mo17 and A7) was obtained from Dr. M. Motto, Instituto Sperimentale per La Cerealicoltura, Bergamo, Italy; inbred lines D102, DK105, D107, D118, D140, D503, D44, D01, D403, D406, CO125 and W401 were obtained from Dr. W.G. Polmer, University of Hohenheim, Stuttgart, Germany; 88 recombinant inbred lines derived from the cross B73xMo17 were obtained from Dr. C.W. Stuber, North Carolina State University, Raleigh, NC, USA; oilseed rape DNA (12 genotypes obtained from Van der have, The Netherlands) were isolated using a modified CTAB procedure described by Stewart and Via (Stewart, C.J., Jr and Via L.E. (1993) *Biotechniques* **14,** 748-750.) All restriction enzymes were purchased from Pharmacia (Pharmacia LKB Biotechnology AB, Uppsala, Sweden), except for the restriction enzyme *MseI*, which was purchased from New England Biolabs Inc. (Beverly, MA, USA). TaqStart™ Antibody was obtained from Clontech (Clontech Laboratories, Palo Alto, CA, USA).

[0139] T4 DNA ligase and T4 polynucleotide kinase were also obtained from Pharmacia. All PCR reagents and consumables were obtained from Perkin Elmer Corp. (Norwalk, CT, USA). Radioactive reagents were purchased from Amersham (Amersham International plc, Little Chalfont, Buckinghamshire, UK) or Isotopchim (Isotopchim SA, Ganagobie, France).

B. AFLP primers and adapters used.

[0140] All oligonucleotides were made on a Biotronic Synostat D DNA-synthesizer (Eppendorf GmbH, Maintal, Germany) or Milligen Expedite 8909 DNA-synthesizer (Millipore Corp. Bedford, MA, USA). The quality of the crude oligonucleotides was checked by end-labeling with polynucleotide kinase and [γ-^{33}P]ATP and subsequent electrophoresis on 18% denaturing polyacrylamide gels (Maxam, A.M. and Gilbert, W. (1980) *Methods Enzymol.* **65,** 499-560.).

[0141] Oligonucleotide adapters and primers for AFLP™ analysis were generally used without further purification when they proved to be >85% full length.

[0142] For the rare cutter site two different AFLP adapters were used: 1) the conventional AFLP adapter (called AFLP adapter) as target site for primer annealing and 2) an AFLP adapter serving as blocking agent (called e.g. *PstI*-adapter*). Both adapters consist of a core sequence (CORE) and a site-specific sequence (SITE) (Sadri, R. and Hornsby, P.J. (1996) *Nucleic Acids Res.* **24,** 5058-5059). The blocking adapters differ from the AFLP adapters only in core sequence. This is illustrated below for *PstI* and *HpaII*-adapters.

	CORE	SITE
*Pst*I-adapter:	5'-CTCGTAGACTGCGTACA	**TGCA**-3'
	3'-CATCTGACGCATGT-5'	
*Hpa*II-adapter:	5'-CTCGTAGACTGCGTACA-3'	
	3'-CATCTGACGCATGT	**GC**-5'
*Pst*I-adapter*:	5'-GCATCAGTGCATGCG	**TGCA**-3'
	3'-GTAGTCACGTACGC-5'	
*Hpa*II-adapter*:	5'-GCATCAGTGCATGCG-3'	
	3'-GTAGTCACGTACGC	**GC**-5'

The conventional and blocking adapter for *Msp*I and *Cla*I are identical to those for *Hpa*II.

[0143] For the frequent cutter site *Mse*I, also two different AFLP-adapters were used: 1) a *Mse*I-adapter only for non-selective amplification (called *Mse*I-adapter$^+$) and 2) a *Mse*I-adapter (called *Mse*I-adapter) as annealing target site for further nonselective and selective amplification. The *Mse*I-adapters differ only in core sequence. This is illustrated below:

	CORE	SITE
*Mse*I-adapter:	5'-GACGATGAGTCCTGAG-3'	
	3'-TACTCAGGACTC	**AT**-5'
*Mse*I-adapter$^+$:	5'-CTCGTAGACTGCGTACC-3'	
	3'-CTGACGCATGG	**AT**-5'

AFLP primers consist of three parts, a core sequence, a site-specific sequence (SITE) and a selective extension (EXT) (vide EP-A-0534858). This is illustrated below for *Pst*I- and HpaII-primers with three selective nucleotides (shown as NNN):

	CORE	SITE	EXT
*Pst*I	5'-GACTGCGTACA	TGCAG	NNN-3'
*Hpa*II	5'-GACTGCGTACA	CGG	NNN-3'
*Cla*I	5'-GACTGCGTACA	CGAT	NNN-3'

AFLP-primers for *Msp*I and *Hpa*II have a similar architecture. However, it should be noted that the AFLP-primers for *Pst*I, *Hpa*II, *Msp*I and *Cla*I are designed only for the AFLP adapters.

[0144] The two *Mse*I-primers are distinguished in the same way as the two *Mse*I-adapters: the MseI-primer has the MseI-adapter as annealing target site, while MseI-primer$^+$ has the MseI-adapter$^+$ as annealing target site. The difference between the two *Mse*I-primers is shown below:

	CORE	SITE	EXT
*Mse*I-primer:	5'-GATGAGTCCTGAG	TAA	NNN-3
*Mse*I-primer$^+$:	5'-GTAGACTGCGTACC	TAA-3'	

C. Modification of DNA and template preparation.

[0145] The protocols A, B and C below describe the generation of A-,B- and C-templates for AFLP reactions using the enzyme combination (EC) *Pst*I/*Mse*I. Protocol C is equivalent to the standard AFLP protocol as described in EP-A-0 534 858 and by Vos et al (Vos, P., Hogers, R., Bleeker, M., Reijans, M., van de Lee, T., Hornes, M., Frijters, A., Pot, J., Peleman, J., Kuiper, M. and Zabeau, M. (1995) *Nucl. Acids Res.* **23**, 4407-4414).

Protocol A:

[0146] Genomic DNA (0.5 µg) is incubated for 1 h at 37°C with 5 U *Pst*I and 5 U *Mse*I in 40 µl 10mM Tris-HAc pH 7.5, 10 mM MgAc, 50 mM KAc, 5 mM DTT, 50 ng/µl BSA (acetylated). Next, 10 µl of a solution is added, containing 5 pMol *Pst*I-adapter$^+$, 50 pMol *Mse*I-adapter, 50 pMol *Mse*I-adapter$^+$, 1U T4 DNA-ligase, 1mM ATP in 10 mM Tris-HAc pH 7.5, 10 mM MgAc, 50 mM KAc, 5 mM DTT, 50 ng/µl BSA, and the incubation is continued for 3 h at 37°C. Adapters are prepared by adding equimolar amounts of both strands; adapters are not phosphorylated. After ligation, the reaction mixture is diluted to 250 µl with 10 mM Tris-HCl, 0.1 mM EDTA pH 8.0, and used for PCR amplification or stored at -20°C.

Protocol B:

[0147] Genomic DNA (0.5 µg) is incubated for 1 h at 37°C with only 5 U *Mse*I in 40 µl 10mM Tris-HAc pH 7.5, 10 mM MgAc, 50 mM KAc, 5 mM DTT, 50 ng/µl BSA (acetylated). Next, 10 µl of a solution is added, containing 50 pMol *Mse*I-adapter, 50 pMol *Mse*I-adapter$^+$, 1U T4 DNA-ligase, 1mM ATP in 10 mM Tris-HAc pH 7.5, 10 mM MgAc, 50 mM KAc, 5 mM DTT, 50 ng/µl BSA, and the incubation is continued for 3 h at 37°C. Adapters are prepared as above (Protocol A). After ligation, the reaction mixture is diluted to 250 µl with 10 mM Tris-HCl, 0.1 mM EDTA pH 8.0, and used for PCR amplification or stored at - 20°C.

Protocol C:

[0148] Genomic DNA (0.5 µg) is incubated for 1 h at 37°C with 5 U *Pst*I and 5 U *Mse*I in 40 µl 10mM Tris-HAc pH 7.5, 10 mM MgAc, 50 mM KAc, 5 mM DTT, 50 ng/µl BSA (acetylated). Next, 10 µl of a solution is added, containing 5 pMol *Pst*I-adapter, 50 pMol *Mse*I-adapter, 1U T4 DNA-ligase, 1mM ATP in 10 mM Tris-HAc pH 7.5, 10 mM MgAc, 50 mM KAc, 5 mM DTT, 50 ng/µl BSA, and the incubation is continued for 3 h at 37°C. Adapters are prepared as above (Protocol A). After ligation, the reaction mixture is diluted to 500 µl with 10 mM Tris-HCl, 0.1 mM EDTA pH 8.0, and used for PCR amplification or stored at - 20°C.

D. Synthesis of unmethylated DNA by nonselective PCR amplification and further modification of DNA and template preparation

[0149] Synthesis of unmethylated A- and B-templates is performed by nonselective PCR amplification. This nonselective PCR amplification is performed with the following cycle profile for 7 cycles: a 30 s DNA denaturation step at

94°C, a 1 min annealing step at 56°C and a 2 min extension step at 72°C. Amplifications are performed in 20 µl containing 5 µl template-DNA, 30 ng *Mse*I-primer, 30 ng *Mse*I-primer+, 0.4 U *Taq* polymerase, 10 mM Tris-HCl pH 8.3, 1.5 mM $MgCl_2$, 50 mM Kcl and 0.2 mM of all four dNTPs.

[0150] After amplification, the reaction mixtures A and B are incubated again for 1 h at 37°C with 5 U *Pst*I in 40 µl 10mM Tris-HAc pH 7.5, 10 mM MgAc, 50 mM KAc, 5 mM DTT, 50 ng/µl BSA(acetylated). Next, 10 µl of a solution is added, containing 5 pMol *Pst*I-adapter, 1U T4 DNA-ligase, 1mM ATP in 10 mM Tris-HAc pH 7.5, 10 mM MgAc, 50 mM KAc, 5 mM DTT, 50 ng/µl BSA(acetylated), and the incubation is continued for 3 h at 37°C.

[0151] After ligation, the reaction mixture is diluted to 1000 µl with 10 mM Tris-HCl, 0.1 mM EDTA pH 8.0, and used for PCR amplification or stored at - 20°C.

[0152] Generation of A- and B-templates for AFLP reactions using methylation-sensitive rare cutters leaving a 5'-extension (e.g. *Hpa*II, *Msp*I, *Cla*I) involves inactivation of the remaining *Taq* polymerase after the nonselective amplification, to avoid incorporation of remaining dNTPs after restriction. This is achieved by adding 220 ng TaqStart™ Antibody/ U *Taq* polymerase to the amplification mixture, prior to further restriction and ligation.

E. AFLP reactions

[0153] Amplification reaction conditions are described using DNA templates (A, B and C) for the EC *Pst*I/*Mse*I. With other ECs, the procedure is identical except for the use of appropiate primers.

[0154] AFLP fingerprinting of large genomes generally involves an amplification in two steps. The first step of this amplification procedure, named pre-amplification, employes two AFLP primers, one aimed at the *Pst*I-ends and one at the *Mse*I-ends, each having a single selective nucleotide. These primers are not radioactively labelled. Amplifications are performed in 20 µl containing 5 µl template-DNA, 30 ng *Mse*I-primer, 30 ng *Pst*I-primer, 0.4 U *Taq* polymerase, 10 mM Tris-HCl pH 8.3, 1.5 mM $MgCl_2$, 50 mM KCl and 0.2 mM of all four dNTPs.

[0155] After pre-amplification, the reaction mixtures are diluted 10-fold with 10 mM Tris-HCl, 0.1 mM EDTA pH=8.0, and used as templates for the second amplification reaction.

[0156] The second amplification reaction again employes two oligonucleotide primers, one aimed at the *Pst*I-ends and one at the *Mse*I-ends, each having two or three selective nucleotides. The *Pst*I-primer is radioactively end-labeled using [γ-^{33}P] ATP and T4 polynucleotide kinase. The labelling reactions are performed in 50 µl 25 mM Tris-HCl pH 7.5, 10 mM $MgCl_2$, 5mM DTT, 0.5mM spermidine-3HCl using 500 ng oligonucleotide primer, 100 µCi [γ-^{33}P] ATP (1000-3000Ci/mol) and 10 U T4 polynucleotide kinase. Amplifications are performed in 20 µl containing 5 µl template-DNA, 5 ng labeled *Pst*I-primer (0.5 µl from the labelling reaction mixture), 30 ng *Mse*I-primer, 0.4 U *Taq* polymerase, 10 mM Tris-HCl pH 8.3, 1.5 mM $MgCl_2$, 50mM KCl and 0.2 mM of all four dNTPs.

[0157] AFLP preamplification reactions are performed for 24 cycles (protocol A and B) and 20 cycles (protocol C) with the following cycle profile: a 30 s DNA denaturation step at 94 °C, a 1 min annealing step at 56°C and a 1 min extension step at 72°C. AFLP reactions with primers having two or three selective nucleotides are performed for 36 cycles with the following cyle profile: a 30 s DNA denaturation step at 94°C, a 30 s annealing step (see below) and a 1 min extension step at 72°C. The annealing temperature in the first cycle is 65 °C, and is subsequently reduced each cycle by 0.7 °C for the next 12 cycles, then continued at 56 °C for the remaining 23 cycles. All amplification reactions are performed in a PE-9600 thermocycler (Perkin Elmer Corp. Norwalk, CT, USA).

F. Gel analysis

[0158] Following amplification, reaction products are mixed with an equal volume (20 µl) of formamide dye (98% formamide, 10 mM EDTA pH 8.0, and bromophenol blue and xylene cyanol as tracking dyes). The resulting mixture is heated for 3 min at 90 °C, then quickly cooled on ice. Each sample (2 µl) was loaded on a 5% denaturing (sequencing) polyacrylamide gel (Maxam, A.M. and Gilbert, W. (1980) *Methods Enzymol.* **65**, 499-560). The gel matrix is prepared using 5% acrylamide, 0.25% methylene bisacryl, 7.5 M urea in 50 mM Tris/50 mM Boric acid/1mM EDTA (pH 8.3). To 100 ml of gel solution 500 µl of 10 % APS and 100 µl TEMED is added and gels are cast using a SequiGen 38x50 cm gel apparatus (Biorad Laboratories Inc., Hercules, CA, USA). 100 mM Tris/100mM Boric acid/2ml EDTA was used as running buffer. Electrophoresis is performed at constant power, 110W, for about 2 hours. After electrophoresis, gels are fixed for 30 min in 10 % acetic acid, rinsed with water for 10 min, dried on the glass plates and exposed to Fujix phosphorimage screens for 16 h. Fingerprint patterns are visualized using a Fujix BAS-2000 phosphorimage analysis system (Fuji Photo Film Company Ltd., Japan).

Example II: Methylation AFLP fingerprinting of large genomes.

[0159] Initial experiments with methylation AFLP fingerprinting of plant genomic DNAs indicated that the complexity of the template libraries is of the same order as in the standard AFLP protocol. For maize, however, the number of the

fragments detected in the B reaction (corresponding to the fragments of the A and C reaction jointly) by *PstI*+2/*MseI*+3 primer combinations (PCs) is substantially elevated, so 'tuning' to a slightly higher selectivity of the PCs is desirable and results in combined *PstI* PCs, e.g. *PstI*+AGW or *PstI*+AGS), where W stands for A and T jointly, and where S stands for C and G jointly.

Figures 7A and 7B show methylation AFLP fingerprints of genomic DNAs from the three maize inbred lines A7, B73 and Mo17. The two panels show *PstI*/*MseI* fingerprints, corresponding with the following PCs (from left to right): I. *PstI*+AGW/*MseI*+CTT, and II. *PstI*+AGS/*MseI*+CTT . Lane A, B and C represent the corresponding A, B and C restriction fragments, referring to the native methylated state of the rare cutter sites. The molecular weight size range of the fingerprints is 200-500 nucleotides.

Theoretically, every band present in lane A or lane C should also be present in lane B, because lane B represents both the native methylated and unmethylated rare cutter sites jointly. As seen in Figure 2, for PCs I, this is true for more than 90% of the fragments in lane A and C. In some cases (e.g. PC II), however, the percentage of fragments in lane A and C present in lane B can be low. An explanation for this is the following: short *MseI*-*MseI* restriction fragments have a competitive advantage over the longer *MseI*-*MseI* restriction fragments in the nonselective amplification. Consequently, rare cutter restriction sites (methylated or not), on average residing more on long *MseI*-*MseI* restriction fragments, will have a selective disadvantage. Subsequent restriction and ligation of the underrepresented fragments will generate a reduced amount of templates.

[0160] Although, as suggested by the fingerprints, large plant genomes consist predominantly of unique AFLP fragments, the presence of repetitive DNA is reflected by fragments of moderate (arrow 1) or high (arrow 2) band intensities in Figures 7A and 7B, in the good accordance with the correlation that exists between the AFLP band intensity and the original number of restriction fragments (Vos, P., Hogers, R., Bleeker, M., Reijans, M., van de Lee, T., Hornes, M., Frijters, A., Pot, J., Peleman, J., Kuiper, M. and Zabeau, M. (1995) *Nucl. Acids Res.* **23,** 4407-4414). Since these multicopy restriction fragments are more abundantly present in lane A than in lane C, it can be concluded that moderate and highly repetitive sequences are strongly methylated. It should also be noted that the excessive use of primers by these multicopy *PstI*-*MseI* restriction fragments also leads to an underrepresentation of fragments in lane B.

Example III: Estimating the extent of ^{5m}CpNpG as presented in *PstI*-sites of some large plant genomes.

[0161] In plants the modified C at position 5 (^{m5}C) is not only confined to CpG dinucleotides as is in animals, but also occurs at a variety of other cytosine containing dinucleotides, all of which are part of the basic trinucleotide CpNpG where N can be any nucleotide (Gruenbaum, Y., Naveh-Maney, T., Cedar, H. and Razin, A., 1981, *Nature* **292,** 860-862) The restriction enzyme *PstI*, having 5'-CTGCAG-3' as recognition site, containing two CpNpG trinucleotides, is sensitive to methylation at the 5'-C and the 3'-C (McClelland, M., Nelson, M. and Raschke, E. (1994) *Nucl. Acids Res.* **22,** 3640-3659). Whether simultaneous methylation of both C's of the *PstI* recognition site is possible, is not clear. A mean percentage ± standard error of methylated *PstI*-sites in the nuclear DNA of tomato, maize and oilseed rape, are given in Table 1.

Table 1

Mean percentages ± standard error of methylated *PstI*-sites and ^{5m}CpNpG as presented in *PstI*-sites in the nuclear DNA of tomato, maize and oilseed rape, and published mean percentages ^{5m}C residues (expressed as % ^{5m}C/^{5m}C +C) in the nuclear DNA of tomato and maize. N= number of genotypes assayed. n = total number of *PstI*/*MseI* restriction fragments counted.

Species	N	% *PstI*-sites methylated	% CpNpG-sites methylated	n	% ^{5m}C/^{5m}C+C
tomato	6	39.75 ± 1.00	19.88 ± 0.50	219 - 248	22.8 (19)
maize	12	50.78 ± 1.33	25.39 ± 0.67	753 - 820	28.9 (20)
oilseed rape	6	33.19 ± 0.62	16.60 ± 0.31	548 - 609	

[0162] It is clear from Table 1 that variation in the percentage of methylated *PstI*-sites is very low within a species. Using the percentage of ^{5m}CpNpG-sites as presented in *PstI*-sites, calculated as half the number of methylated *PstI*-sites, as estimation of the percentage of ^{5m}C residues in the nuclear DNA (expressed as ^{5m}C/^{5m}C +C), results in percentages ^{5m}C residues very similar to published percentages, based on HPLC analysis (Messuguer, R., Ganal, M.W., Steffens, J.C., and Tanksley, S.D., 1991, *Plant Mol.Biol.* **16,** 753-770 and Montero, L.M. et al., 1992, *Nucl.Acids Res.* **20,** 3207-3210), for tomato and maize (see Table 1). Therefore, methylation of only one C residue in the *PstI* recognition

site is highly probable.

Example IV: Estimating the extent of ^{5m}CpG as presented in *Hpa*II-, *Msp*I- and *Cla*I-sites in some large plant genomes.

[0163] *Msp*I and *Hpa*II (methylation isoschizomers) have the same recognition site 5'-CCGG-3', containing one CpG dinucleotide. However, *Msp*I is sensitive to 5'-^{5m}C, whereas *Hpa*II is sensitive to methylation at position 5 of either C (McClelland, M., Nelson, M. and Raschke, E. (1994) *Nucl. Acids Res.* **22**, 3640-3659). Therefore, *Msp*I and *Hpa*II are appropriate for estimating the extent of ^{5m}CpG. The extent of ^{5m}CpG methylation by using *Hpa*II and *Msp*I as methylation-sensitive rare cutters is measured as the difference in the number of bands counted in a *Msp*I and *a Hpa*II fingerprint in lane C; the difference in the number of bands counted in a *Msp*I and a *Hpa*II fingerprint in lane A must give the same result. The percentages of ^{5m}CpG sites in the nuclear DNA of tomato and maize, as presented in *Hpa*II- and *Msp*I-sites, are measured only for one genotype/species, and are given in Table 2.

[0164] Another methylation-sensitive restriction enzyme *Cla*I, having the CpG-dinucleotide containing recognition site 5'-ATCGAT-3', and affected by ^{5m}CpG (McClelland, M., Nelson, M. and Raschke, E. and (1994) *Nucl. Acids Res.* **22**, 3640-3659) was also found to be suitable as rare cutter in the detection of ^{5m}CPG. The percentages of ^{5m}CpG as presented in *Hpa*II-, *Msp*II- and *Cla*I-sites in the nuclear DNA of tomato, maize and oilseed rape, are measured for only one genotype/species, and are given in Table 2.

Table 2

Percentages of ^{5m}CpG in the nuclear DNA of tomato, maize and oilseed rape, as presented in *Hpa*II-, *Msp*I- and *Cla*I-sites. n = total number of restriction fragments counted.

Species	^{5m}CpG as presented in *Cla*I-sites	n	^{5m}CpG as presented in *HpaII/Msp*I-sites	n
tomato (cv.52201)	57.9	2833	57.6	2169
maize (B73)	53.9	2694	39.6	1791
oilseed rape (T528)	48.4	2973		

[0165] In tomato, a similar percentage of ^{5m}CpG as presented in *Cla*I-sites is obtained as presented in either *Hpa*II- or *Msp*I-sites, in contrast with maize. Estimations of the percentage of ^{5m}C residues as presented in *Cla*I-sites in the nuclear DNA of tomato, mais and oilseed rape is significantly higher than the the corresponding percentages ^{5m}C residues as presented in *Pst*I-sites. This can suggest that the majority of ^{5m}C residues in the tomato, maize and oilseed rape genome exist at CpG sites.

Example V: ^{5m}C polymorphism and its inheritance.

[0166] AFLP primarily detects variation of the primary DNA sequence, either base substitutions or DNA rearrangements. Nucleotides modified by methylation are not considered to be a part of the primary nucleotide sequence of an individual. This nucleotide modification, resulting from a post-replicative event at defined but not all target sequences (i.e. CpG and CpNpG), represents two additional forms of DNA polymorphism: (i) polymorphism reflecting the variation in the primary nucleotide sequence of the methylated restriction site and/or variation in the restriction size (^{m}AFLP markers), and (ii) allele-specific methylation (^{asm}AFLP markers). Since DNA-methylation is also the only source of allelic difference between epialleles, ^{asm}AFLP markers are epiallelic markers. Since methylation AFLP provides a positive display of the native methylated sites, it allows exploitation of this additional sources of DNA polymorphism.

For methylation polymorphism to be useful in e.g. AFLP mapping studies, it must be stably inherited. Figure 8A shows the segregation of a number of ^{m}AFLP markers in a Recombinant Inbred (RI) population derived from the cross B73 x Mo17. The band intensities segregate into two distinct classes, homozygous absent and homozygous present, approximating the Mendelian 1:1 segregation. To identify possible ^{asm}AFLP marker pairs, the following criteria are applied: (i) an AFLP band and a ^{m}AFLP band might be epialleles when derived from different parents, with the same PC; (ii) the AFLP and ^{m}AFLP marker map to the same locus (complementary segregation). Figures 8A and 8B show the segregation of a pair of ^{asm}AFLP markers in the RI population. The band intensities segregate into two distinct classes, homozygous absent and homozygous present. Of a total of 673 mapped AFLP + ^{m}AFLP markers, generated by a set of 14 *PstI/MseI* PCs, ^{m}AFLP markers account for 44,6%, which is in good accordance with the 50.78% methylated *Pst*I-sites in the maize genome (Table 1), while hardly 1% of the marker alleles behave like epialleles.

As can be seen from the above description and Examples, the method of invention can be considered a DNA

fingerprinting technique that detects genomic restriction fragments and resembles in that respect the AFLP technique, with the major difference that methylation AFLP displays native methylated sites too, but in general only if they are located in the recognition site of the obligate methylation-sensitive rare cutter. As for AFLP, the multiplex ratio is high (50-100 restriction fragments) and is a function of (i) the cleavage frequency of the methylation-sensitive rare cutter enzyme and (ii) the number and nature of the selective bases of the specific primer set.

[0167] Also, as demonstrated in the Examples, the method of the invention can be used for estimating the extent of ^{5m}CpNpG, resp. ^{5m}CpG, as presented in *Pst*I-, resp. *Cla*I-, *Hpa*I- and *Msp*I-sites in the genome of few crops. Although generalization of these estimates to percentages ^{5m}CpNpG and ^{5m}CpG in the genome, irrespective the recognition site, may in some cases not be fully representative as methylation of C can be biased by the nature of its flanking nucleotides (recognition site), or by the percentage G+C of the genomic region it resides (CG-islands), the method of the invention will still provide valuable information to the artisan.

[0168] Although hereinabove the invention was particularly described with respect to the use of methylation AFLP to estimate the extent of ^{5m}CpG and ^{5m}CpNpG as presented in some recognition sites in the genome, the invention and its possible uses are not only limited to that. Modifications of C at other positions than 5 (e.g. ^{m4}C), or modified nucleotides other than C (e.g. ^{m6}A), whether present in a symmetrical or a non-symmetrical dinucleotide (e.g. ^{5m}CpC), are reported ((Gruenbaum, Y., Naveh-Maney, T., Cedar, H. and Razin, A., 1981, *Nature* **292,** 860-862); Selker, E.U., Fritz, D.Y. and Singer, M.J., 1993, *Science* **262,** 1724-1728; and Meyer, P., Niedenhof, I., ten Lohuis, M., 1994, *The EMBO Journal* **13,** 2084-2088).

[0169] Methylation AFLP can also be used with advantage to estimate the low prevalence of these minor nucleotide modifications presented by recognition sites of methylation-sensitive rare cutters, and to shed a light on their distributions over the genome.

[0170] Methylation AFLP allows the exploitation of two additional forms of DNA polymorphism: (i) polymorphism reflecting variation in the primary nucleotide sequence of the methylated restriction site and/or variation in the restriction fragment size, and (ii) allele-specific methylation. Hardly 1% of the marker alleles behave like epialleles.

[0171] By monitoring the segregation of ^{m}AFLP and ^{asm}AFLP markers in a RI population, it was shown that (i) C-methylation segregates in perfect accordance with the primary target sequence, while (ii) the few epialleles observed in the parents inherit in a Mendelian fashion to the offspring.

[0172] Like AFLP markers, most ^{m}AFLP markers correspond to unique positions in the genome, and, hence, can be exploited as landmarks in and as bridging tools between genetic and physical maps (Vos, P., Hogers, R., Bleeker, M., Reijans, M., van de Lee, T., Hornes, M., Frijters, A., Pot, J., Peleman, J., Kuiper, M. and Zabeau, M. (1995) *Nucl. Acids Res.* **23,** 4407-4414). Therefore, methylation AFLP may be useful in genomic research. Native methylated sites are present on cloned DNA segments, e.g., yeast artificial chromosomes (YACs) as unmethylated sites. Hence, native methylated sites can not be distinguished from native unmethylated sites on a physical map. However, ligning up the physical map with a genetic map containing ^{m}AFLP markers, helps to identify native methylated sites on the physical map.

[0173] More than in plant genetics, in human genetics the epigenetic effects of DNA methylation at CpG-island containing promotors is well documented. Due to its sensitive, reliable and quantitative nature, methylation AFLP is an attractive technique to determine the DNA methylation levels at specific gene loci like tumor-suppressor genes and to trace imprinted genes.

Note: AFLP™ is a trademark of Keygene N.V.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:
(A) NAME: Keygene N.V.
(B) STREET: Agro Business park 90
(C) CITY: Wageningen
(D) STATE: Gelderland
(E) COUNTRY: The Netherlands
(F) POSTAL CODE (ZIP): NL-6700 AE

(ii) TITLE OF INVENTION: Method for detecting nucleic acid methylation using AFLPTM

(iii) NUMBER OF SEQUENCES: 15

(iv) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)

(2) INFORMATION FOR SEQ ID NO: 1:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 nucleotides
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(vi) ORIGINAL SOURCE: Synthetic

(ix) FEATURES:
(D) OTHER INFORMATION: PstI-adapter

(ix) FEATURES:
(A) NAME/KEY: misc-feature
(B) LOCATION: 4...17
(D) OTHER INFORMATION: complementary to SEQ ID NO:2

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

CTCGTAGACT GCGTACATGC A 21

(2) INFORMATION FOR SEQ ID NO: 2:
(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 14 nucleotides
(B) TYPE: nucleic acid (C) STRANDEDNESS: double
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(vi) ORIGINAL SOURCE: Synthetic

(ix) FEATURES:
(D) OTHER INFORMATION: PstI-adapter

(ix) FEATURES:

(A) NAME/KEY: misc-feature
(B) LOCATION: 1...14
(D) OTHER INFORMATION: complementary to SEQ ID NO:1

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

TGTACGCAGT CTAC 14

2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 nucleotides
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(vi) ORIGINAL SOURCE: Synthetic

(ix) FEATURES:
(D) OTHER INFORMATION: HpaII-adapter

(ix) FEATURES:
(A) NAME/KEY: misc-feature
(B) LOCATION: 4...17
(D) OTHER INFORMATION: complementary to SEQ ID NO:4

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3
CTCGTAGACT GCGTACA 17

2) INFORMATION FOR SEQ ID NO: 4: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 16 nucleotides (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(vi) ORIGINAL SOURCE: Synthetic; HpaII-adapter

(ix) FEATURES:
(D) OTHER INFORMATION: HpaII-adapter

(ix) FEATURES:
(A) NAME/KEY: misc-feature
(B) LOCATION: 2...16
(D) OTHER INFORMATION: complementary to SEQ ID NO:3

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

CGTGTACGCA GTCTAC 16
(2) INFORMATION FOR SEQ ID NO: 5: (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 19 nucleotides
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(vi) ORIGINAL SOURCE: Synthetic

(ix) FEATURES:
(D) OTHER INFORMATION: PstI-adapter

(ix) FEATURES:
(A) NAME/KEY: misc-feature
(B) LOCATION: 2...15
(D) OTHER INFORMATION: complementary to SEQ ID NO:6

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

GCATCAGTGC ATGCGTGCA 19

(2) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 14 nucleotides
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(vi) ORIGINAL SOURCE: Synthetic

(ix) FEATURES:
(D) OTHER INFORMATION: PstI-adapter

(ix) FEATURES:
(A) NAME/KEY: misc-feature
(B) LOCATION: 1...14
(D) OTHER INFORMATION: complementary to SEQ ID NO:5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

CGCATGCACT GATG 14

(2) INFORMATION FOR SEQ ID NO: 7: (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 15 nucleotides
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(vi) ORIGINAL SOURCE: Synthetic

(ix) FEATURES:
(D) OTHER INFORMATION: HpaII-adapter

(ix) FEATURES:
(A) NAME/KEY: misc-feature
(B) LOCATION: 2-15
(D) OTHER INFORMATION: complementary to SEQ ID NO:8

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

GCATCAGTGC ATGCG 15

(2) INFORMATION FOR SEQ ID NO: 8:

(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 16 nucleotides (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(vi) ORIGINAL SOURCE: Synthetic

(ix) FEATURES:
(D) OTHER INFORMATION: HpaII-adapter

(ix) FEATURES:
(A) NAME/KEY: misc-feature
(B) LOCATION: 2-16
(D) OTHER INFORMATION: complementary to SEQ ID NO:7

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

CGCGCATGCA CTGATG 16

(2) INFORMATION FOR SEQ ID NO: 9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 15 nucleotides
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(vi) ORIGINAL SOURCE: Synthetic

(ix) FEATURES:
(D) OTHER INFORMATION: HpaII-adapter

(ix) FEATURES:
(A) NAME/KEY: misc-feature
(B) LOCATION: 5-16
(D) OTHER INFORMATION: complementary to SEQ ID NO:10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

GCATCAGTGC ATGCG 15

(2) INFORMATION FOR SEQ ID NO: 10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 16 nucleotides
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(vi) ORIGINAL SOURCE: Synthetic

(ix) FEATURES:
(D) OTHER INFORMATION: HpaII-adapter

(ix) FEATURES:
(A) NAME/KEY: misc-feature
(B) LOCATION: 2-16
(D) OTHER INFORMATION: complementary to SEQ ID NO:9

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

CGCGCATGCA CTGATG 16

(2) INFORMATION FOR SEQ ID NO: 11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 nucleotides
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(vi) ORIGINAL SOURCE: Synthetic

(ix) FEATURES:
(D) OTHER INFORMATION: MseI-adapter

(ix) FEATURES:
(A) NAME/KEY: misc-feature
(B) LOCATION: 6...17
(D) OTHER INFORMATION: complementary to SEQ ID NO:12

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

CTCGTAGACT GCGTACC 17

(2) INFORMATION FOR SEQ ID NO: 12

(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 13 nucleotides (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: unknown (ii) MOLECULE TYPE: DNA

(vi) ORIGINAL SOURCE: Synthetic

(ix) FEATURES:
(D) OTHER INFORMATION: MseI-adapter

(ix) FEATURES:
(A) NAME/KEY: misc-feature
(B) LOCATION: 3-13
(D) OTHER INFORMATION: complementary to SEQ ID NO:11

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

TAGGTACGCA GTC 13

(2) INFORMATION FOR SEQ ID NO: 13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 16 nucleotides
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(vi) ORIGINAL SOURCE: Synthetic

(ix) FEATURES:
(D) OTHER INFORMATION: PstI-primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

GACTGCGTAC ATGCAG 16

(2) INFORMATION FOR SEQ ID NO: 14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 14 nucleotides
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(vi) ORIGINAL SOURCE: Synthetic

(ix) FEATURES:
(D) OTHER INFORMATION: ClaI-primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

GACTGCGTAC ACGG 14

(2) INFORMATION FOR SEQ ID NO: 15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 15 nucleotides
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(vi) ORIGINAL SOURCE: Synthetic

(ix) FEATURES:
(D) OTHER INFORMATION: ClaI-primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:

GACTGCGTAC ACGAT 15

Claims

1. Use of AFLP in analysing or determining the state of methylation or of the methylation pattern of a starting DNA

and/or in distinguishing between methylated and non-methylated sites in a starting DNA.

2. Method for analysing or determining the methylation pattern of a starting DNA and/or for distinguishing between methylated and non-methylated sites in a starting DNA, comprising at least

(A) generating a first DNA fingerprint, containing bands corresponding to both the methylated and non-methylated sites of interest; and/or
(B) generating a second DNA fingerprint, containing bands corresponding only to the methylated sites of interest;
and optionally comprising
(C) generating a third DNA fingerprint, containing bands corresponding only to the non-methylated sites of interest;
and optionally further comprising
(D) analysing the fingerprint(s) thus obtained.

3. Method accoring to claim 2, in which at least two, and preferably all three of the fingerprints resulting from (A), (B) and/or (C) are generated and compared to each other.

4. Method according to claim 2 or 3, in which the respective fingerprints resulting from (A), (B) and/or (C) are generated using an amplification technique, preferably AFLP.

5. Method according to any of claims 2-4, in which the fingerprints resulting from (A), (B) and/or (C) are generated in conjunction with each other, preferably from the same starting DNA, DNA-pool or DNA-preparation.

6. Method according to any of claims 2-5, in which the fingerprints resulting from (A), (B) and/or (C) are run in seperate lanes of the same electrophoresis gel.

7. Method for producing a fingerprint containing bands corresponding to both the methylated sites in a starting DNA, in particular for use in the method of claims 2-5, comprising:

1. restricting the starting DNA with a frequent cutter and a methylation sensitive rare cutter;
2. ligating the restriction fragments thus obtained to a first and second adapter for the frequent cutter, as well as a first adapter for the rare cutter;
3. amplifying the mixture thus obtained, using primers for the first and second frequent cutter adapter;
4. restricting the amplified fragments thus obtained using a methylation sensitive rare cutter;
5. ligating the restriction fragments thus obtained to a second adapter for the methylation sensitive rare cutter;
6. amplifying the restriction fragments thus obtained using a primer for the first or second frequent cutter adapter, and a primer for the second rare cutter adapter;
7. optionally analysing the mixture of amplified fragments thus obtained in a manner known per se, preferably by generating a DNA fingerprint.

8. Method for producing a fingerprint containing bands corresponding to both the methylated and the non-methylated sites in a starting DNA, in particular for use in the method of claims 2-5, comprising:

1. restricting the starting DNA with a frequent cutter;
2. ligating the restriction fragments thus obtained with to third and fourth adapter for the frequent cutter;
3. amplifying the mixture thus obtained, using primers for the third and fourth adapter for the frequent cutter;
4. restricting the amplified fragments thus obtained using a methylation-sensitive rare cutter;
5. ligating the restriction fragments thus obtained to a third adapter for the methylation sensitive rare cutter;
6. amplifying the mixture thus obtained, using a primer for the third methylation sensitive rare cutter adapter, as well a primer for the third or fourth frequent cutter adapter;
7. optionally analysing the mixture of amplified fragments thus obtained in a manner known per se, preferably by generating a DNA fingerprint.

9. Method according to claim 7 or 8, in which the third and fourth frequent cutter adapter are the essentially same as the first and second frequent cutter adapter.

10. Method according to claim 7, 8 or 9, in which the third rare cutter adapter is essentially the same as the second rare

cutter adapter.

11. Method according to any of claims 7-9, in which the frequent cutter restriction enzyme is chosen from *Mse*I and *Taq*I and the methylation sensitive rare cutter is chosen from *Pst*I, *Hpa*II, *Msp*I, *Cla*I, *Hha*I, *EcoR*II, *Bbv*I, *Pvu*II, *Xma*I, *Sma*I, *Nci*I, *Ava*I, *Hae*II, *Sal*I, *Xho*I and *Pvu*II.

12. Method according to any of claims 2-6 or 7-11 for use in distinguishing, identifying or classifying individuals, varieties or species on the basis of their DNA-methylation, by comparing methylation patterns of different individual sources; for fingerprinting of genomes and detecting methylation polymorphisms; or for detecting specific methylation patterns corresponding to the presence of specific sequences or genetic traits, as well as inherited DNA-methylation and allele-specific methylation.

13. Method according to any of claims 2-6 or 7-11 for use in distinguishing transformed and not-transformed genes or sequences, for instance in the detection of transgenes; and distinguishing replicated and non-replicated DNA in bacteria.

14. Method according to any of claims 2-6 or 7-11 for use in investigating or determining the regulation of gene expression in eukaryotes; in particular for detecting gene silencing or gene activation.

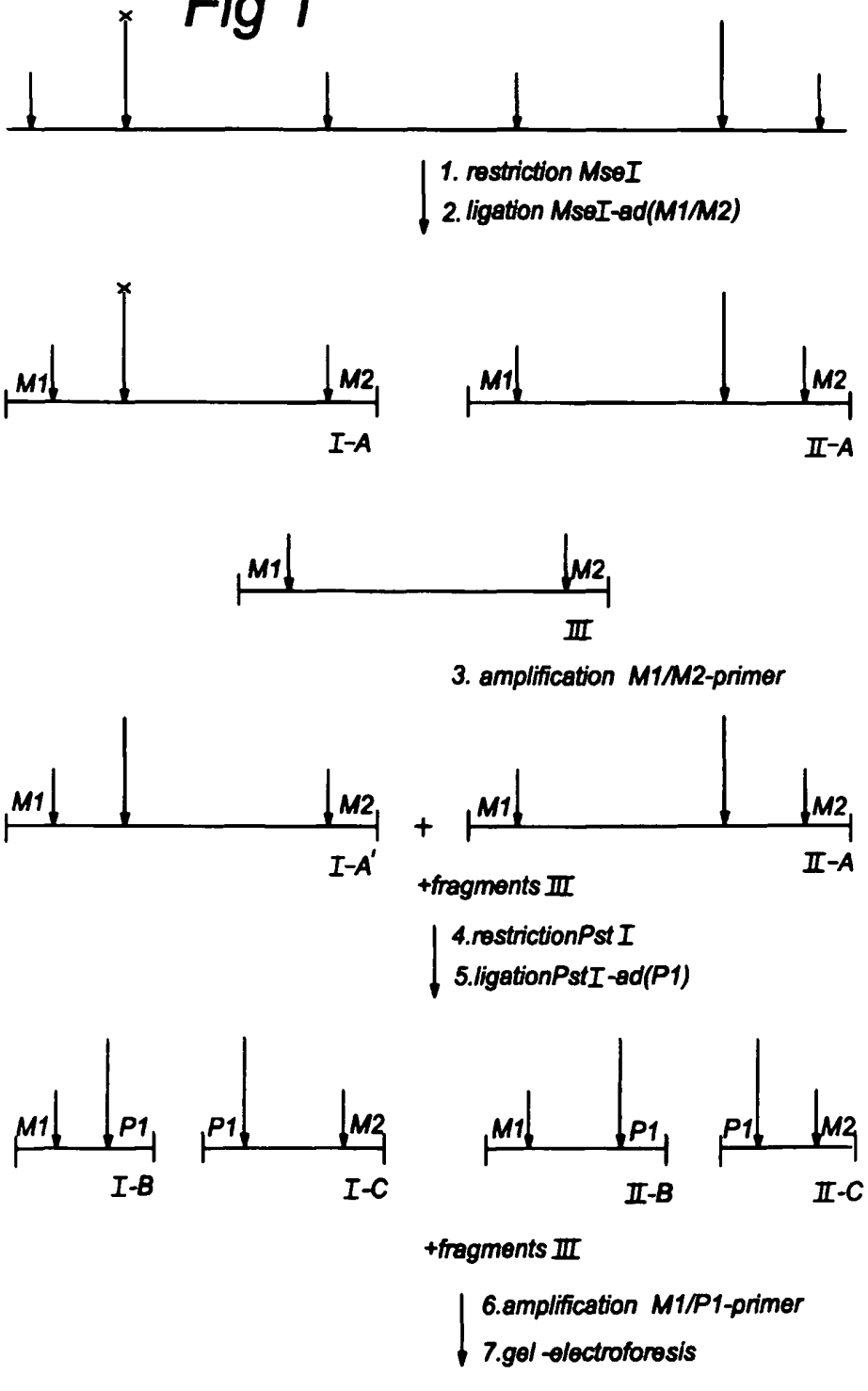
15. Method according to any of claims 2-6 or 7-11 for use in detecting of genetic disorders, such as detecting DNA methylation associated with genetic diseases and "imprinting", as well as increased susceptibility for mutagenesis and cancer;

16. Method according to any of claims 2-6 or 7-11 for use in detecting of gen for estimating the extend of methylation in a DNA or genome; or for detecting methylated DNA-markers.

17. Kit for use with the method of any of claims 1-16, said kit comprising at least a frequent cutter restriction enzyme; a methylation sensitive rare cutter restriction enzyme; (at least) a first and second adapter for use with the frequent cutter; and (at least) a first and second adapter for use with the rare cutter; as well as primers for use with these adapters; and optionally further components for kits known per se.

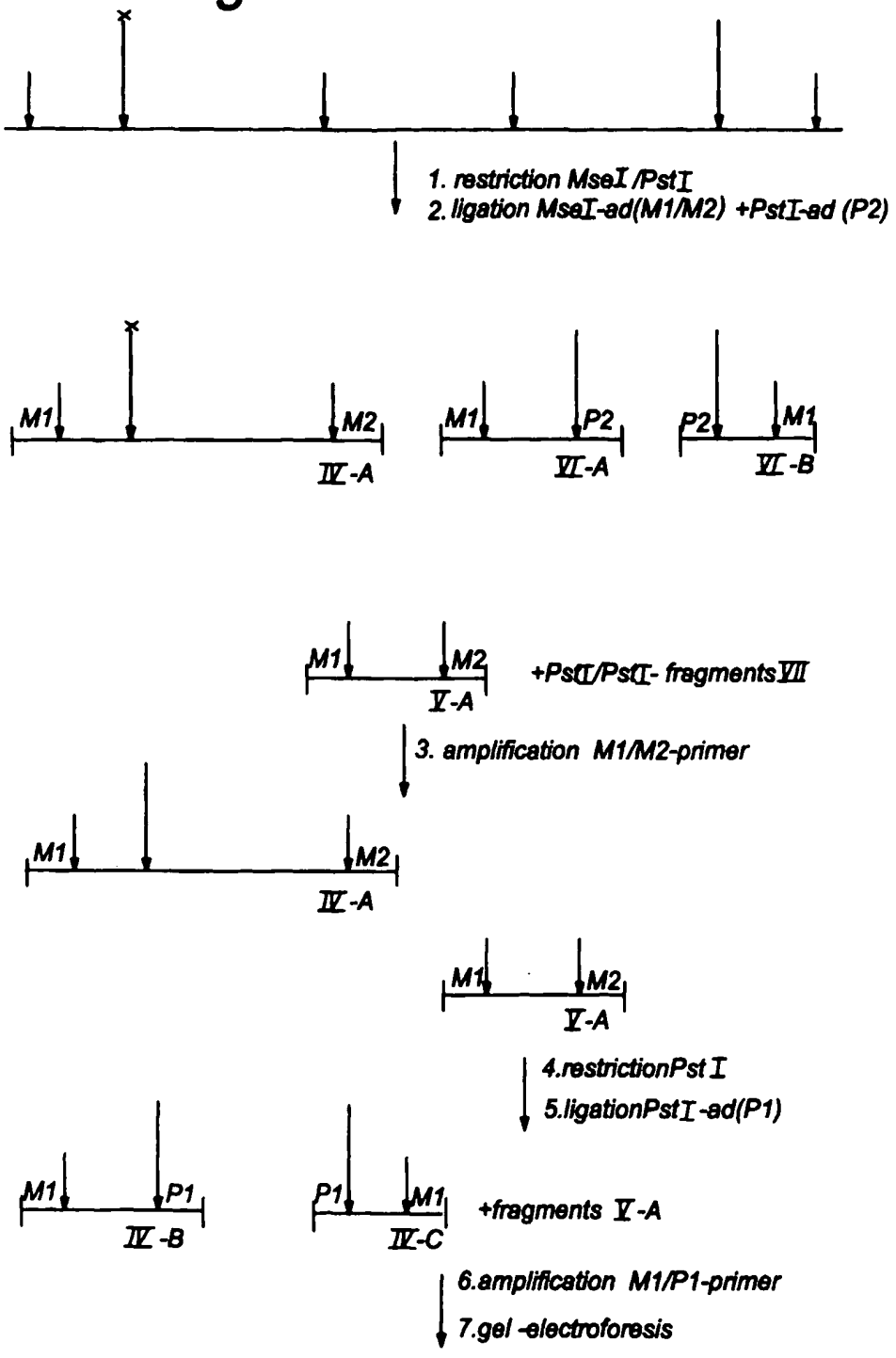
18. Data generated by the method of any of claims 1-16, optionally on a suitable data carrier, including DNA-fingerprints, photographs or other reproductions thereof, as well as (stored) analog or digital data thereon.

Fig 1



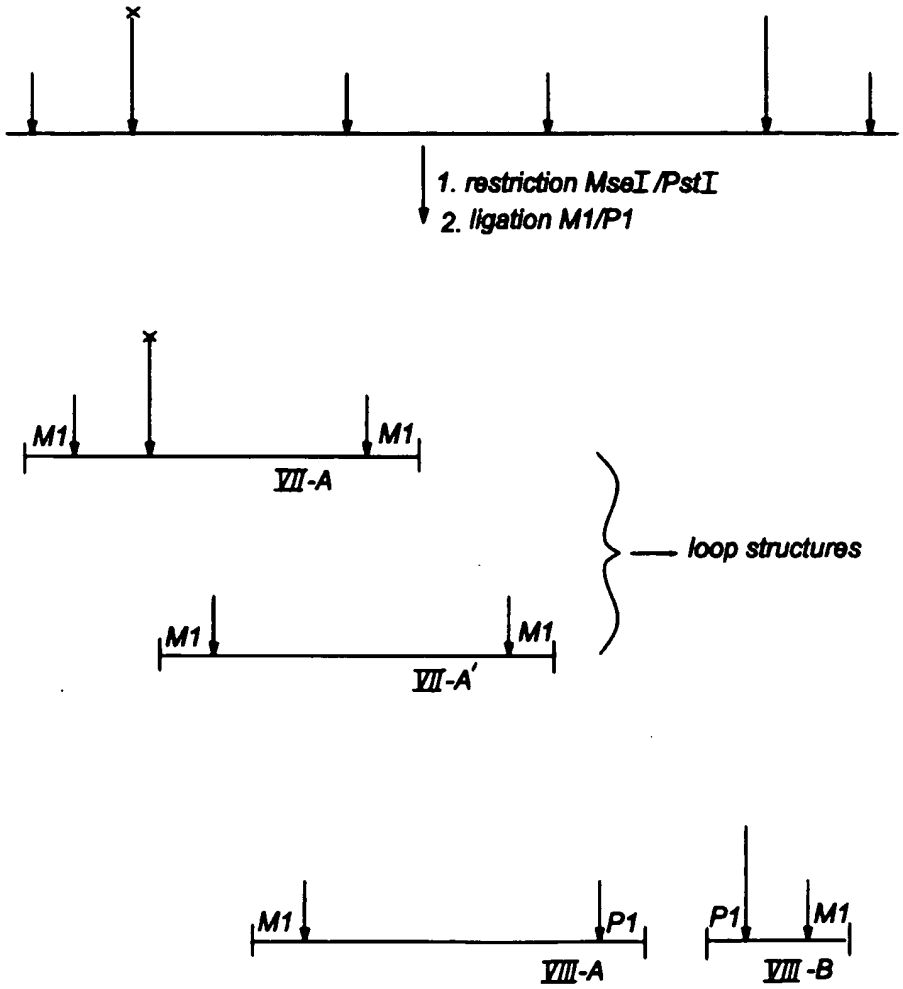
bands corresponding to I-B , I-C , II-B , II-C .

Fig 2



bands corresponding to IV-B and IV-C

Fig 3



bands corresponding to VIII-A and VIII-B

Fig 4

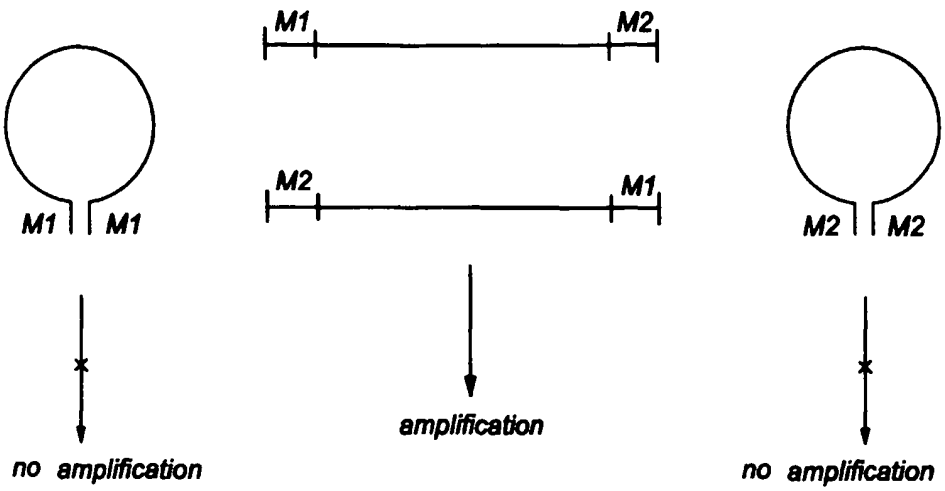

M1
M2
M2
M1
M1
M1
M2
M2
amplification
no amplification
no amplification

Fig 5

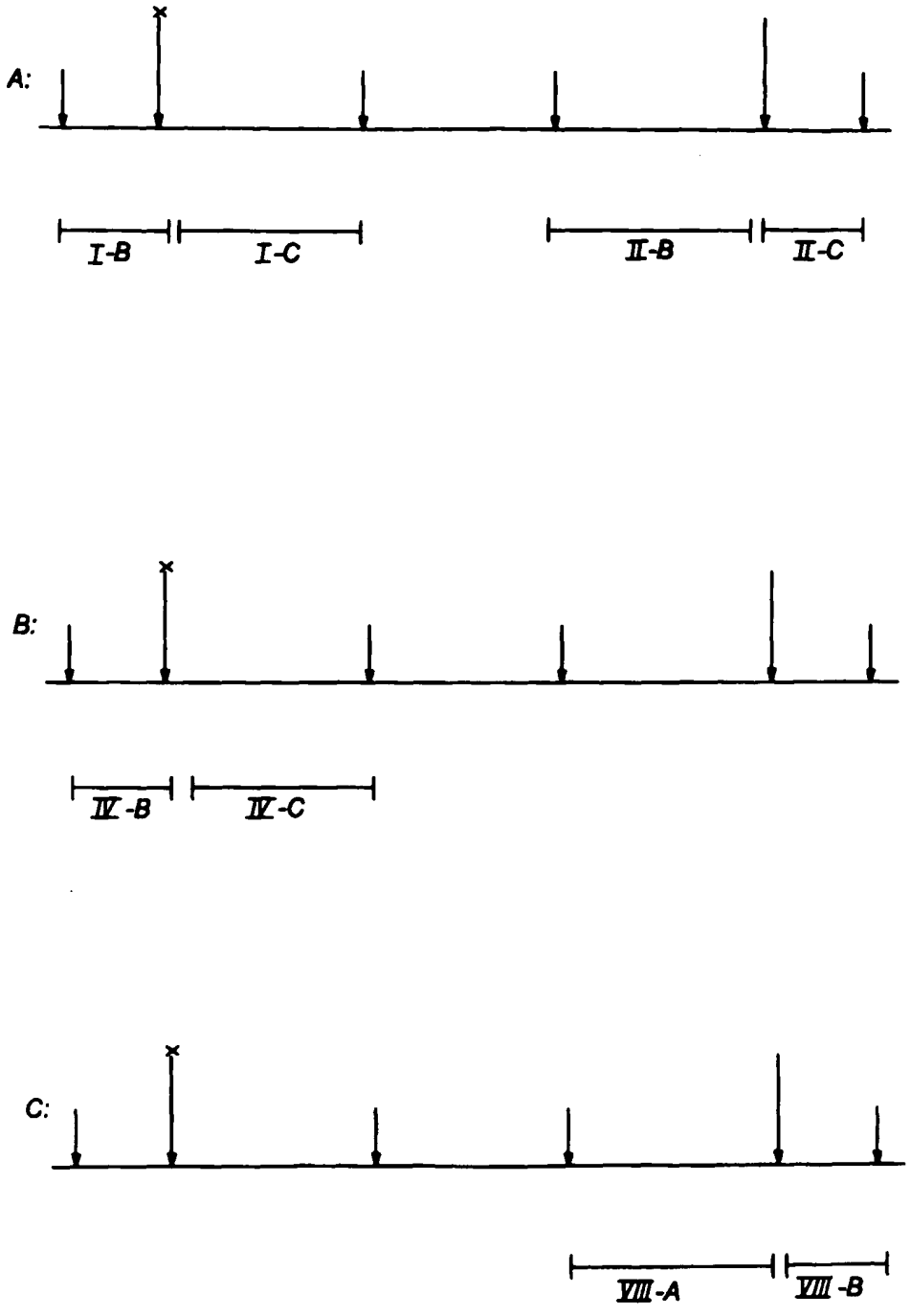


Fig 6

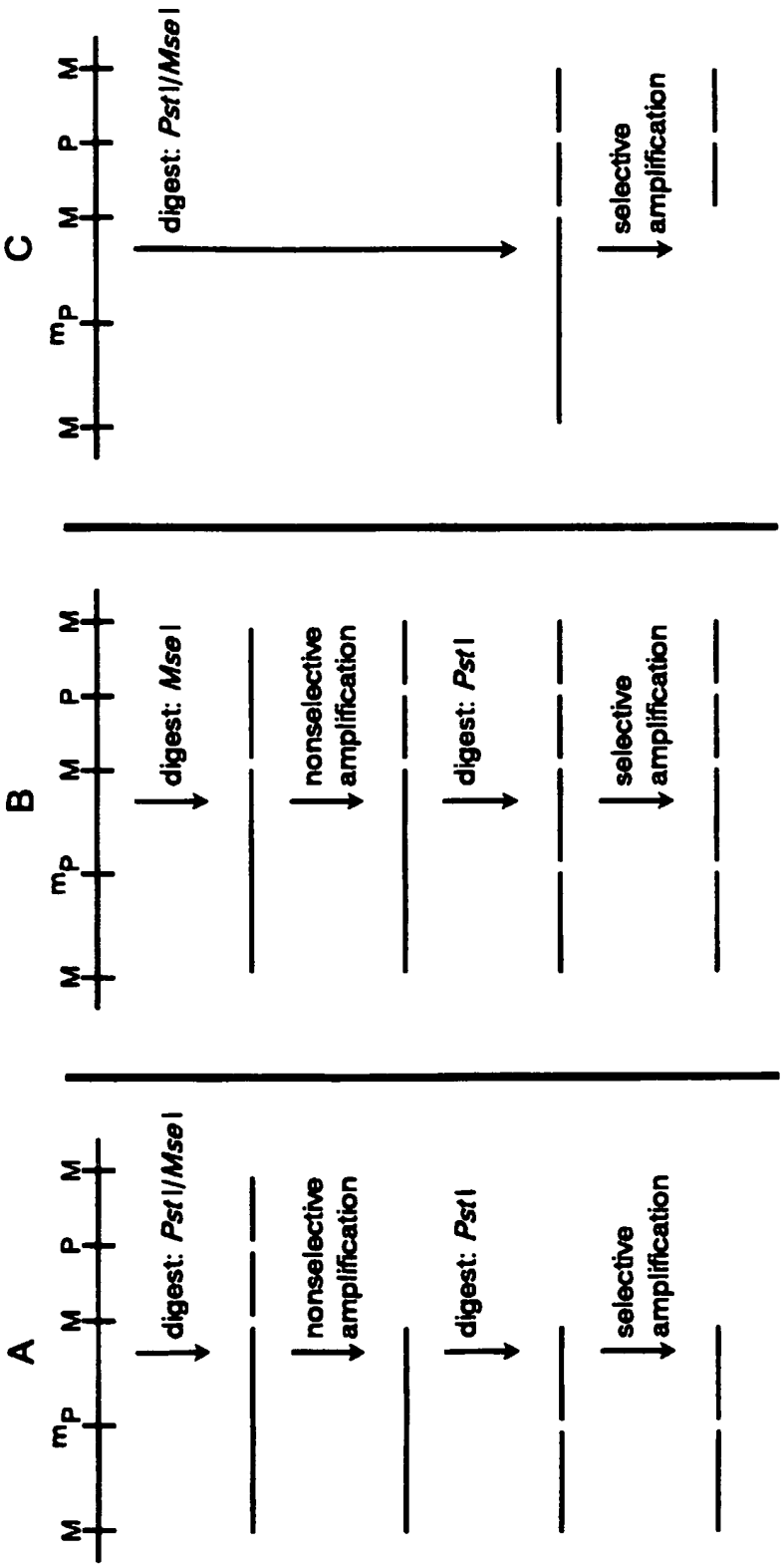

A
M
mP
M
P
M
digest: PstI/MseI
nonselective amplification
digest: PstI
selective amplification
B
M
mP
M
P
M
digest: MseI
nonselective amplification
digest: PstI
selective amplification
C
M
mP
M
P
M
digest: PstI/MseI
selective amplification

Fig 7a

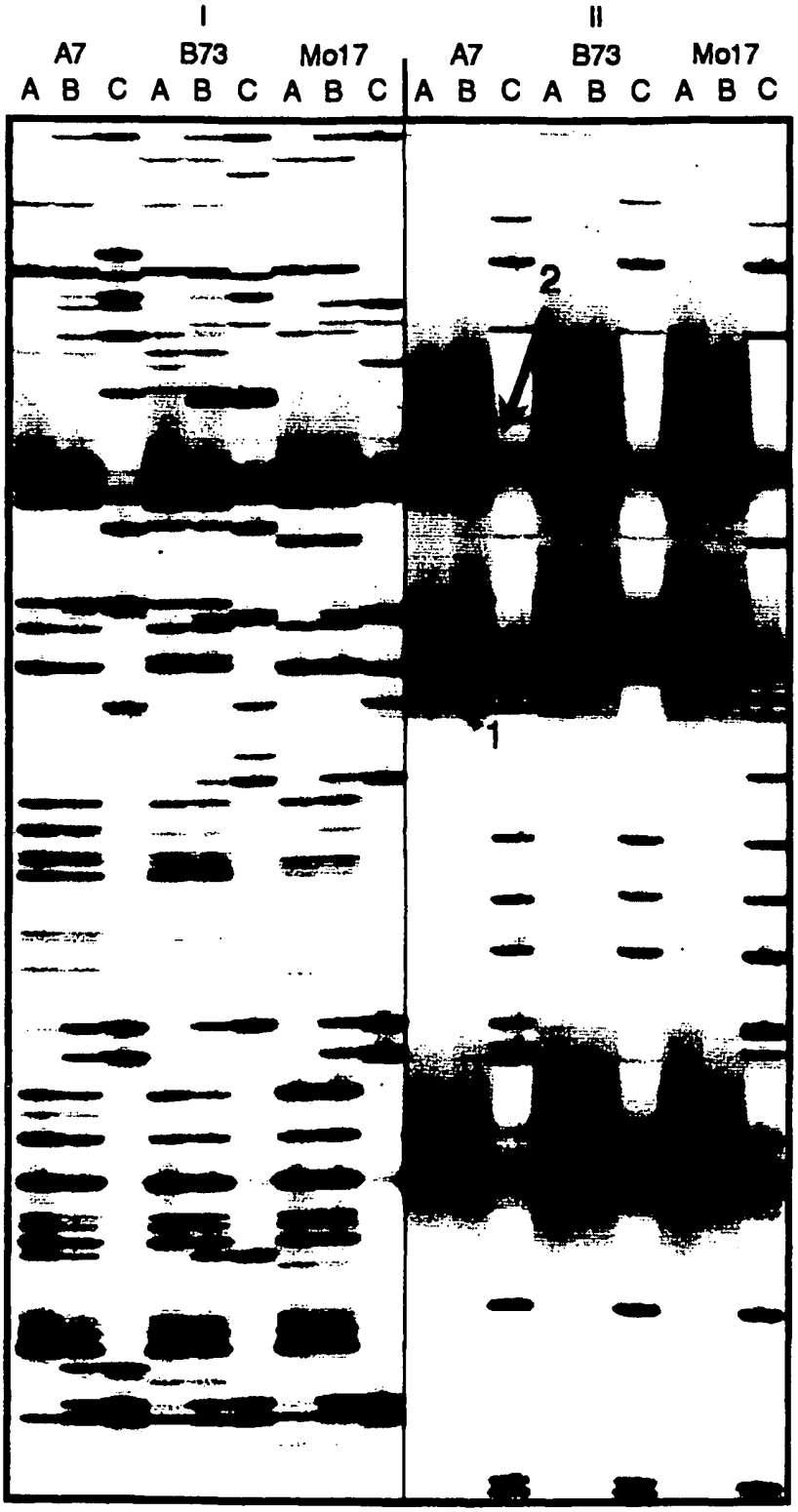
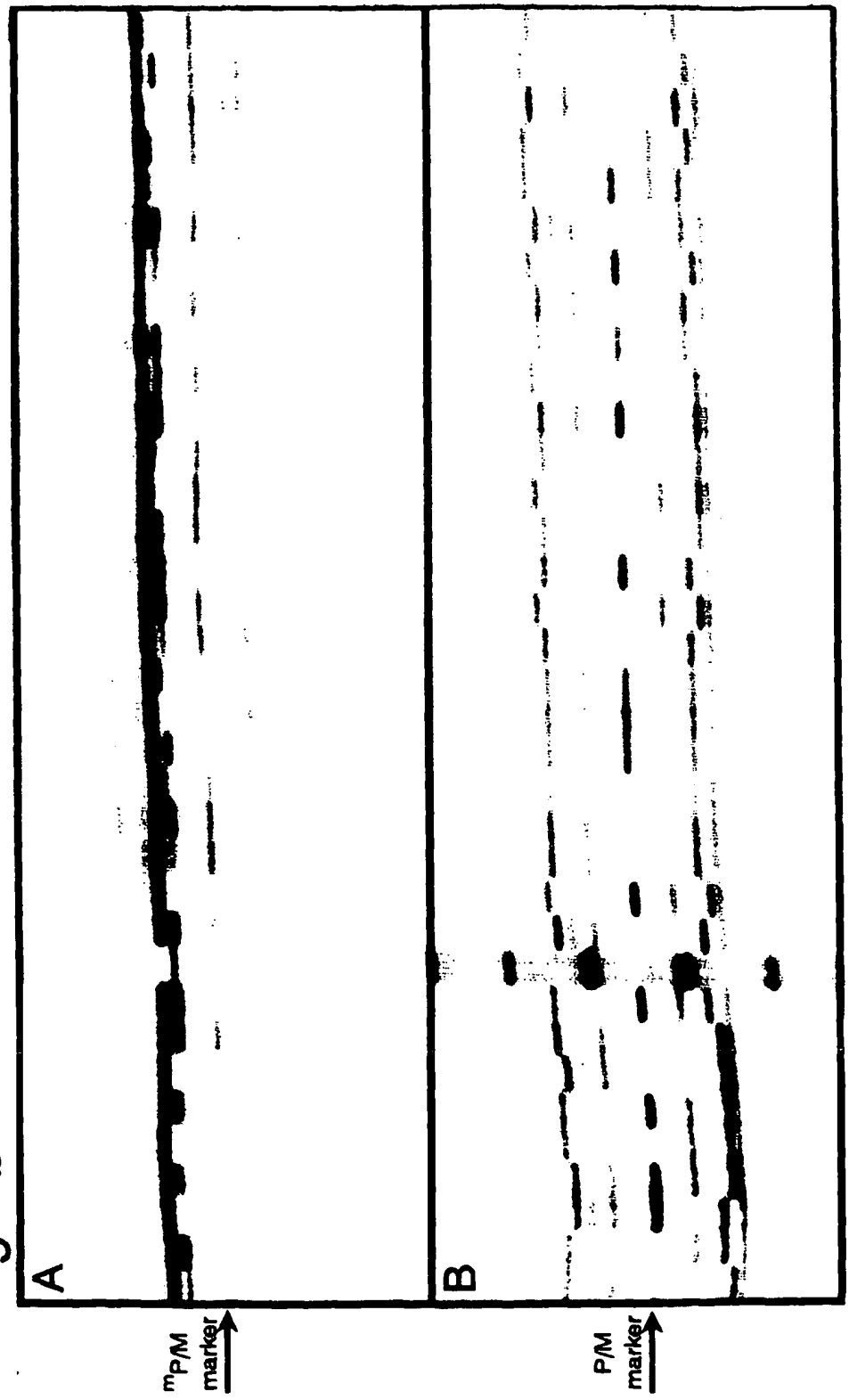


Fig 7b



European Patent Office

EUROPEAN SEARCH REPORT

Application Number
EP 98 20 2549

DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION
X	BIOSIS:AN:1997:439076; Donini P. et al.,Genome 40,p.521-526,1997	1,16,17	C12Q1/68
nti	ng reveals pattern differences between template DNA extracted from different plant organs XP002110060 * the whole document * ---		
Y	Medline:AN1998120884 Chen Z. et al.,Brit.J.Cancer 77, 181-185,1998 Methylation of CpG island is not a ubiquitous mechanism for the loss of oestrogen receptor in breast cancer cells XP002110075 * the whole document * ---	1-18	
Y,D	EP 0 534 858 A (KEYGENE NV) 31 March 1993 * the whole document * ---	1-18	
Y	WO 97 45560 A (NORTH SHORE UNIVERSITY HOSPITA) 4 December 1997 * the whole document * ---	1-18	TECHNICAL FIELDS SEARCHED C12Q
Y,D	VOS P. ET AL.,: "AFLP: a new technique for DNA fingerprinting" NUCLEIC ACIDS RESEARCH, vol. 23, no. 21, - 1995 pages 4407-4414, XP002109691 * the whole document * ---	1-18	
D,A	WO 96 27024 A (AVITECH DIAGNOSTICS INC) 6 September 1996 * the whole document * ---		
A	WO 98 08981 A (LIFE TECHNOLOGIES INC) 5 March 1998 * the whole document * ---		
	-/--		

The present search report has been drawn up for all claims

Place of search	Date of completion of the search	Examiner
THE HAGUE	23 July 1999	Müller, F

CATEGORY OF CITED DOCUMENTS

X : particularly relevant if taken alone
Y : particularly relevant if combined with another document of the same category
A : technological background
O : non-written disclosure
P : intermediate document

T : theory or principle underlying the invention
E : earlier patent document, but published on, or after the filing date
D : document cited in the application
L : document cited for other reasons

& : member of the same patent family, corresponding document

EPO FORM 1503 03.82 (P04C01)

European Patent Office

EUROPEAN SEARCH REPORT

Application Number
EP 98 20 2549

DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION
P,X	XIONG LZ. ET AL.,: "Patterns of cytosine methylation in an elite rice hybrid and its parental lines, detected by a methylation-sensitive amplification polymorphism technique" MOL. GEN. GENET., vol. 261, - April 1999 pages 439-446, XP002110074 * the whole document *	1-18	

TECHNICAL FIELDS SEARCHED

The present search report has been drawn up for all claims

Place of search	Date of completion of the search	Examiner
THE HAGUE	23 July 1999	Müller, F

CATEGORY OF CITED DOCUMENTS

X : particularly relevant if taken alone
Y : particularly relevant if combined with another document of the same category
A : technological background
O : non-written disclosure
P : intermediate document

T : theory or principle underlying the invention
E : earlier patent document, but published on, or after the filing date
D : document cited in the application
L : document cited for other reasons

& : member of the same patent family, corresponding document

EPO FORM 1503 03.82 (P04C01)

ANNEX TO THE EUROPEAN SEARCH REPORT ON EUROPEAN PATENT APPLICATION NO. EP 98 20 2549

This annex lists the patent family members relating to the patent documents cited in the above-mentioned European search report. The members are as contained in the European Patent Office EDP file on
The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

23-07-1999

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
EP 0534858 A	31-03-93	AU 672760 B	17-10-96
		AU 2662992 A	27-04-93
		CZ 9400669 A	15-12-94
		WO 9306239 A	01-04-93
		FI 941360 A	24-05-94
		HU 68504 A	28-06-95
		JP 6510668 T	01-12-94
		NO 941064 A	20-05-94
		ZA 9207323 A	30-08-93
WO 9745560 A	04-12-97	US 5871917 A	16-02-99
		AU 3228597 A	05-01-98
WO 9627024 A	06-09-96	US 5843649 A	01-12-98
		AU 5171496 A	18-09-96
WO 9808981 A	05-03-98	AU 4174397 A	19-03-98

EPO FORM P0459

For more details about this annex : see Official Journal of the European Patent Office, No. 12/82